Stockholm School of Economics
Department of Accounting
Master's Thesis in Accounting and Financial Management
Spring 2017

Predicting Financial Distress

A Model for the European Football Industry

Christoffer Gerdin[λ] **Christian Rump**[φ]

Abstract

The present study successfully develops a probit model that predicts financial distress for European football clubs. The model consists of both financial variables that are common in accounting-based bankruptcy prediction models as well as financial and non-financial variables that capture the distinct characteristics of the football industry. While it significantly outperforms naïve decision rules in classifying clubs as distressed or surviving, it achieves a lower accuracy compared to those attained in many prominent bankruptcy prediction studies. The lower predictive accuracy can be explained by the unique conditions under which football clubs operate. The model is based on a sample of 208 European clubs in the 2006-2016 period that mostly are private or public limited companies and that play in UEFA's top divisions. Using the model at hand, stakeholders of football clubs can make better-informed and more timely decisions in their interactions with clubs.

Keywords: *Bankruptcy prediction; Financial ratio analysis; Probit regression; Football; Soccer*

Tutor: Kenth Skogsvik, Professor, Department of Accounting and Financial Management

Date: May 22nd, 2017

[λ] 40846@student.hhs.se
[φ] 40868@student.hhs.se

Acknowledgements

We would like to extend our sincerest gratitude to our tutor Kenth Skogsvik for invaluable input during the entire thesis-writing and modelling process. Furthermore, we are grateful to Per-Olov Edlund for guidance on statistical matters and to Martin Carlsson-Wall for sharing your sports-business expertise at the early stages of our work.

Many thanks also to Piotr Bukanski, Tiago Enes, Jan Havelka and Stefano Pagliani for your language support in Polish, Portuguese, Czech and Italian, respectively. Without you, the process of writing this thesis would have been far more cumbersome.

Table of Contents

1. Introduction....4

2. Literature Review....6

2.1 Financial Situation in European Football....6
2.2 Predicting Financial Distress....10
2.3 Bankruptcy Prediction Models in Football Literature....14
2.4 Research Question....16

3. Methodology....18

3.1 Designing a Model for the Football Industry....18
3.2 Assessing the Model's Quality....23
3.3 Real-world Application....26

4. Data....27

4.1 Definition of Financial Distress....27
4.2 Data Sample....29

5. Empirics and Analysis....33

5.1 Descriptive Statistics....33
5.2 Model Determination....39
5.3 Accuracies....45
5.4 Two-year Model....47

6. Additional Tests....52

6.1 Application of Zmijewski's (1984) Model....52
6.2 Calibration....54

7. Concluding Remarks....54

7.1 Limitations....55
7.2 Suggestions for Future Research....57

References....58

Appendices....63

1. Introduction

"...running as normal companies, the top leagues in Spain, England and Italy would be bankrupt within two years." - Hembert, Lucero, Mesnard and Rothenbücher (2010, p. 1).

The quote, stated in A.T. Kearney's sustainability study conducted on European football in 2010, gives worrying insight into the financial situation of football clubs in Europe but it also hints at the uniqueness of the industry. Even though an improvement in overall operating and bottom-line income has been observed in recent years, European football clubs still make aggregated net losses of €323 million (UEFA, 2017). Due to financial reasons, a large number of clubs have struggled to survive and have faced difficulties in retaining league positions. In 2015, Parma Calcio[1] was declared bankrupt and was relegated from the Italian first division to the fourth division and, consequently, to amateur football. Similarly, Glasgow Rangers[2] in Scotland entered insolvency proceedings in February 2012 and was eventually relegated to Scotland's fourth division. These are just two famous examples of occurrences that unfortunately have become commonplace in the European football industry.

The downfall of a football club entails severe financial consequences for the club's shareholders and creditors. However, the effects are more widespread than so. Researchers have argued that cities, even entire regions, are affected since football clubs represent symbols of nationalistic pride (Ascari & Gagnepain, 2006) and that the survival of football clubs perhaps is even more desirable than is survival for companies in other industries (Beech, Horsman & Magraw, 2010). The economic and societal importance of football clubs, coupled with the grave financial situation many clubs are in, present a unique setting for application of models developed to predict financial distress[3]. The ability to predict financial distress for football clubs in advance, and thus being able to intervene to prevent distress from occurring, should be an intriguing notion to the stakeholders of clubs. However, studies that have attempted this application found that prominent bankruptcy prediction models were inappropriate for the football industry, exhibiting low predictive accuracies (Barajas & Rodríguez, 2010; Gerritsen, 2015). The special characteristics of football clubs, together with the exceptional environment they

[1] Parma Football Club Spa (Parma F.C. S.P.A).
[2] RFC 2012 P.L.C.
[3] Commonly referred to as 'bankruptcy prediction' models.

operate in, hinder the application of bankruptcy prediction models that were developed on samples and assumptions of 'regular firms'[4].

Attempts have been made to develop models that are designed specifically for the football industry (Barajas & Rodríguez, 2010; Scelles et al., 2016; Szymanski, 2012). However, these studies did not attempt to predict financial distress in football clubs but rather aimed to understand *why* distress occurs. Furthermore, due to high data requirements and limitations to a single country, these models are not appropriate in practice. The need for an elaborate model that predicts financial distress specifically for European football clubs and that is easily applicable remains. Consequently, the authors of the present study aim to answer the following research question:

Is it possible to develop a model that uses publicly available information to predict financial distress for European football clubs?

A financial distress prediction model of this kind would not only be beneficial for shareholders and creditors of clubs in making investment decisions, but the application of such a model can be useful to numerous other stakeholders. These include, among others, municipalities and local businesses that are heavily dependent on a club, current and potential sponsors, and fans with a genuine interest in the continuity of their club. In addition, the model can be useful for national and European football associations, including UEFA[5], that are concerned about interruptions of the competition in their leagues due to clubs suffering from financial problems. Hence, a financial distress prediction model developed specifically for European football clubs is expected to have a high practical applicability for football club stakeholders, and has been called for by academics (e.g. Barajas & Rodríguez, 2010; Barajas & Rodríguez, 2014; Gerritsen, 2015).

To be able to answer the research question, data on football clubs is required. Financial information on clubs is retrieved from Bureau van Dijk's Amadeus database (2017). If available, annual reports are accessed through registrars of companies in Germany and England or downloaded directly from the webpages of football clubs. In addition, non-financial performance data and club characteristics are

[4] 'Regular firms' refer to companies that are operating in industries that have commonly been used for financial distress prediction models and that have a classical value-maximizing objective, e.g. the manufacturing industry.
[5] The Union of European Football Associations, or UEFA, is the governing body of European football.

retrieved from Europe's largest football database, transfermarkt.com (2017), from fifaindex.com (2017), and Eurostat (2015).

The study is designed as following. It commences with a detailed examination of the unique financial situation of European football clubs. The study proceeds with an analysis of existing bankruptcy prediction literature to understand which statistical methods, industry foci and financial indicators have been used in prominent prediction models. The literature review concludes by addressing studies that have attempted to understand financial distress in the context of European football, after which the research question of the present study is derived. In the third section, the methodology of this study is outlined, and the statistical model as well as the variables included in the model are discussed. The fourth section presents the definition of 'financial distress' that is used in this study and the data collection process. In the fifth section, the results of the development and application of the prediction models are discussed by examining the one-year model and two-year model separately as well as comparing the accuracy to another study in the sixth section. The study concludes with a discussion on the findings, the limitations of the present study and suggestions for future research.

2. Literature Review

The first subsection in this section gives the reader an introduction to the current financial situation in European football. Thereafter, a discussion on bankruptcy prediction literature follows, which gives an overview of the central studies that have been conducted in the research area and which methodology these studies have employed. Next, research that has attempted to apply or develop accounting-based bankruptcy prediction models in the context of football is given extra attention before the section concludes by deriving the research question.

2.1 Financial Situation in European Football

European football has long been criticized for being financially unhealthy and several researchers have observed indicators suggesting that structural weaknesses permeate the industry (Ascari & Gagnepain, 2006; Barajas & Rodríguez, 2010; Barajas & Rodríguez, 2014; Boeri & Severgnini, 2012; Boscá, Liern, Martínez & Sala, 2008; Dietl & Franck, 2007; Georgievski & Zeger, 2016; Mourao, 2012). These weaknesses are mainly related to:

- The revenue structure of clubs, which heavily relies on broadcasting income,
- The high proportion of fixed costs in the cost structure, which mostly is related to player wages,
- The judicial consequences of poor financial situations, namely bailouts and insolvency proceedings.

The revenue structure is similar in European football clubs, where the single largest source of revenue comes from selling broadcasting rights (Boeri & Severgnini, 2012; UEFA, 2017). This type of revenue has furthermore been the fastest-growing[6] in European football and, as such, clubs are very dependent on this source of income (Ascari & Gagnepain, 2006; Boeri & Severgnini, 2012; Georgievski & Zeger, 2016; Morrow & Stephen, 2014; Mourao, 2012; Szymanski, 2012). Due to worse-than-expected performance, especially when a club gets relegated, the revenue from broadcasting rights drops significantly, making it hard for the club to cover the fixed costs that were taken on in expectation of a certain revenue level (Beech, Horsman & Magraw, 2010).

Despite the increase in overall revenues (UEFA, 2017), clubs have not seen higher profits as the revenue growth has been approximately corresponding to increases in player wages[7] and transfer fees (Ascari & Gagnepain, 2006; Hembert et al., 2010). Mourao (2012) argued that clubs' increases in costs actually have *exceeded* the growth in revenues. The increases in costs can be understood by examining the cost structure of clubs, where player wages and amortization entailed by high player transfer fees represent the most significant items (Boscá et al., 2008; Hembert et al., 2010). Players, which are treated as intangible assets on the clubs' balance sheet when they are acquired, constitute the bulk of the clubs' assets (Ascari & Gagnepain, 2006)

Kuper and Szymanski (2014) described the clubs' spending on players as an "arms race" (p. 69). Superior, more expensive players are expected to achieve better outcomes on the football field, which is paramount for clubs in order to avoid relegation. Hence, clubs tend to overspend on player wages to keep up with their peers (Ibid.). Boeri & Severgnini (2012) observed that clubs pay significant premiums for the best players but that these premiums are not matched by an equal increase in revenues. Consequently, it is not uncommon for football clubs to be loss-making (Ascari & Gagnepain,

[6] European football revenues have exhibited a compound annual growth rate of 9.3% between 1996 and 2015. In 2015, 34% of revenues came from broadcasting, making it the single largest source of revenues. (UEFA, 2017).
[7] Player wages have grown at a compound annual growth rate of 10.3% between 1996 and 2015 (UEFA, 2017).

2006; Barajas & Rodríguez, 2014; Boeri & Severgnini, 2012; Kuper & Szymanski, 2014; Szymanski, 2012). In Spain, 88.6% of first and second division clubs made operational losses in the 2007-08 season (Barajas & Rodríguez, 2010). Similarly, 88 out of 107 top-division Italian clubs made net losses in the season ending 2011 (Boeri & Severgnini, 2012). On an aggregate level, European football's net losses were with €1,670 million at their lowest point in recent history in 2011 (UEFA, 2017).

To a large extent, the poor financial condition of clubs originates from unsustainable business models (Hembert et al., 2010; Barajas & Rodríguez, 2013). In particular, clubs have a high proportion of fixed costs, primarily associated with player wages, that are difficult to cover in the short-term if revenues drop due to, for example, relegation or worse-than-expected performance (Georgievski & Zeger, 2016; Mourao, 2012; Szymanski, 2012). Clubs generally respond in one of two ways when such a drop occurs. Either, the club sells assets, e.g. players, to cover the losses and consequently gives up future economic and on-field performance. Alternatively, and often preferably from the club's perspective as there is no impact on the on-field performance, the club commits to elevated levels of short-term debt (Barajas & Rodríguez, 2014; Boscá et al., 2008; Mourao, 2012). Ultimately, if no access to additional capital is found and the situation does not improve, the club faces liquidity issues (Marcinkowska, 2013).

One commonly forwarded reason for the current financial situation in the football industry is that clubs are win-maximizing organizations rather than profit-maximizing[8] (Ascari & Gagnepain, 2006; Barajas & Rodríguez, 2010; Garcia-del-Barrio & Szymanski, 2009; Georgievski & Zeger, 2016; Kuper & Szymanski, 2014). Football clubs aim to achieve the highest amount of on-field success possible, i.e. strive for the best ranking in the league, only limited by the amount of money available to them (Ascari & Gagnepain, 2006; Garcia-del-Barrio & Szymanski, 2009). Kuper and Szymanski (2014) argued that making profits is not the primary purpose of a football club since most stakeholders, even owners, are interested in win-maximization. In fact, Kuper and Szymanski even went as far as to state that success and profit are mutually exclusive for football clubs. Similarly, Hembert et al. (2010) found no correlation between bottom-line financial and on-field performance. Consequently, investors in football clubs often do not expect a positive return on their investment coming directly from the club (Beech et al., 2010; Georgievski & Zeger, 2016; Hembert et al., 2010). These particular investors

[8] "Profit-maximizing" is used interchangeably with "value-maximizing" as the concept commonly is referred to in accounting and finance literature. In the football literature, "profit-maximizing" has become the norm and will therefore be used throughout this study.

represent benefactor-owners that do not have profit-motives (Beech et al., 2010). They can sometimes make profits by attaining an improved business situation in another company that they are simultaneously involved in through reputational spillover or networking benefits from owning the football club, but rarely directly from the investment in the football club (Garcia-del-Barrio, 2009; Hembert et al., 2010; Kuper & Szymanski, 2014).

An additional factor that contributes to the critical financial situation in European football is the reluctance of creditors to demand that football clubs enter into insolvency proceedings and subsequently liquidation (Barajas & Rodríguez, 2010; Cooper & Joyce, 2013; Geogievski & Zeger, 2016; Hembert et al., 2010; Kuper & Szymanski, 2014). The associated reputational risk that creditors face by demanding that a club enters insolvency proceedings or liquidation can severely affect the business operations of the creditor (Cooper & Joyce, 2013). Thus, excessive risk-taking that is irrational for 'regular' companies is not necessarily as irrational for football clubs, knowing that consequences that such actions entail are lenient or even non-occurring.

However, even when insolvency does occur, actual winding-up of clubs is uncommon. Beech et al. (2010) observed that dissolution of clubs happened very rarely in England and they argued that continuity of football clubs was more desirable than continuity of regular companies due to the perceived social relevance of clubs. Szymanski (2010) found that 75 out of 88 clubs that played in the top four divisions in England in 1923 still played in one of these divisions in the 2007-08 season. When football clubs suffer from financial problems they have the possibility to sell players in order to reduce employee costs, which consequently weakens the on-field performance and allows the club to continue business in a lower division (Ibid.). This decrease of product quality is an option regular firms do not benefit from (Ibid.). The longevity and resilience of football clubs is also observed in Spain where clubs enjoy strong local support because they are seen as symbols of nationalistic pride (Ascari & Gagnepain, 2006). Gerritsen (2015) suggested similar reasons for clubs' survival in Dutch football and Hembert et al. (2010) observed such patterns throughout the Big Five[9] leagues. Thus, even if clubs enter insolvency proceedings, dissolutions are rare due to the clubs' relevance to certain members of society and the reluctance of creditors to demand liquidation (Barajas & Rodríguez, 2010; Cooper & Joyce, 2013; Hembert et al., 2010; Szymanski, 2010; Szymanski, 2012).

[9] The biggest football leagues in Europe in terms of revenue, namely: England, France, Germany, Italy and Spain (Hembert et al., 2010).

Given that the dissolution of clubs is unlikely, *voluntarily* entering insolvency proceedings can be viewed as an option for clubs to write off debt and return to a more stable financial situation (Beech et al., 2010; Kuper & Szymanski, 2014; Szymanski, 2010). Insolvency proceedings are common among football clubs in Europe. In England, 66 clubs went through insolvency proceedings between 1982 and 2010 (Szymanski, 2010). Similarly, 22 Spanish clubs entered insolvency proceedings between 2003 and 2011 and 51.4% of Spanish first or second division clubs were close to bankruptcy in the 2007-08 season (Barajas & Rodríguez, 2010; Barajas & Rodríguez, 2013). Furthermore, there were 79 insolvencies in French professional football in the 1970-2014 period (Scelles et al., 2016). Finally, between 2002 and 2011 nine clubs of the Italian first division made use of insolvency proceedings (Boeri & Severgnini, 2012).

The special characteristics inherent in the football industry, combined with the high frequency of clubs entering insolvency proceedings, make European football an interesting area of application for bankruptcy prediction models. To examine the possibility of such an application, the following subsection elaborates on the research that has been done in the bankruptcy prediction literature.

2.2 Predicting Financial Distress

Business failure entails severe consequences for all company stakeholders and the idea of being able to predict failure in advance has long intrigued researchers. Modern financial distress prediction literature has its roots in financial ratio analysis research from the first half of the 20th century. Early studies utilized univariate analysis where one financial ratio at the time was considered and the ratios of failed firms were compared to those of surviving firms (Bellovary, Giacomino & Akers, 2007). However, it was not until Beaver's (1966) study that the accuracy with which financial ratios could predict business failure was tested for the first time (Ibid.).

Beaver (1966) used a paired-sample design where each of the failed firms was matched with a surviving firm with similar characteristics with regard to industry and size to control for differences in these two dimensions. The statistical method employed by Beaver was univariate discriminant analysis where one financial ratio at the time is examined. While Beaver found that the prediction accuracy using a single financial ratio was high when applied to a hold-out sample[10], it was suggested that the accuracy could be improved by utilizing multivariate discriminant analysis (‘MDA’), which allows for considering several ratios at once. The first researcher to use MDA to predict business failure was Altman (1968). Altman’s multivariate model, the so-called Z-score model, uses five financial ratios to classify firms as distressed or surviving and remains one of the most influential in the research area (Bellovary et al., 2007; Grice & Ingram, 2001; Wu, Gaunt & Gray, 2010).

Multivariate discriminant analysis remained the statistical method of choice for researchers during the 1970’s, but during the 1980’s probabilistic prediction models, e.g. logit/probit analysis, emerged and by the 1990’s these types of models had become more prominent than MDA (Bellovary et al., 2007). Whereas MDA results in dichotomous classifications, probabilistic models produce probabilities of failure. For practical application and decision contexts, dichotomous classification models are less useful than models generating a likelihood of failure (Skogsvik & Skogsvik, 2013; Zavgren, 1985). Given that the probability of failure can be incorporated into investment contexts, investors can make better-informed decisions. Furthermore, MDA is based on the assumptions that the independent variables are normally distributed and that the variance-covariance matrices of the predictors are equal for both groups of failed and surviving firms, which several researchers found were rarely, if ever, met (cf. Lennox, 1999; Ohlson, 1980; Skogsvik, 1990; Zavgren, 1985). Ohlson (1980) was one of the first researchers to acknowledge the advantages of probabilistic prediction models and to use such analysis to predict business failure (Bellovary et al., 2007). However, other authors such as Zavgren (1985) argued that while the choice of logit analysis was justified, Ohlson’s model suffered from shortcomings related to the lack of a matched-pairs approach in the sampling and the absence of a hold-out sample for calculating the error rates of the model. Matching firms based on asset size and industry would have controlled for inexplicit factors (Zavgren, 1985), while the use of a hold-out sample is appropriate to validate the accuracy of the model (e.g. Beaver. 1966; Bellovary et al., 2007; Skogsvik, 1987; Zavgren, 1985).

Thus, there is a lack of consensus in the bankruptcy prediction literature and methodological differences between studies are commonplace. Bellovary et al. (2007) compiled the research that has

[10] A hold-out sample represents a sample of firms that was not used for the estimation of the model. It can either consist of a completely new sample of firms or a partition of the estimation sample so that there is no overlap between the firms used for the estimation of the model and those on which the predictive accuracy is tested (Skogsvik, 1987).

been done in the field of bankruptcy prediction since Beaver's (1966) study and concluded that a plethora of different methods have been employed. Not only do the statistical methods differ, but some studies focus on specific industries whereas others are unfocused and consider all industries (Bellovary et al., 2007). The distinction between industries is important to consider given that different financial ratios should be meaningful in different industries (Ibid.). Generally expressed, the institutional context[11] in which the sample firms examined in a bankruptcy prediction model are operating defines application population of the model. For successful application, the institutional context of the firms on which the model is applied should be similar to the institutional context of the sample of firms used to develop the model (Barajas & Rodríguez, 2014; Gerritsen, 2015; Grice & Ingram, 2001).

Table 1 contains a selection of prominent bankruptcy prediction models and offers a comparison of examined countries, industry foci, statistical methods and prediction accuracies. The models listed are by no means the only relevant models in the research area[12]. Rather, *Table 1* contains models found to be pertinent by the authors of the present study. Prediction accuracy is derived from the arithmetic average of the misclassification rates of failed and surviving firms (see *Subsection 3.2.2*). If the accuracy was tested on both the sample of analysis as well as on a hold-out sample, *Table 1* lists the results for the hold-out sample.

Although differences in prediction accuracies exist between the studies listed in *Table 1*, all studies have generated accuracies exceeding those that would have been expected by employing naïve decision rules such as randomly classifying firms based on an a priori probability of distress or classifying all firms according to the outcome with the highest a priori probability of being correct. A second observation is that the industry foci of the studies differ. Whereas some studies have been unfocused, many of the earlier studies were focused on manufacturing or industrial firms. Kim and Gu's (2006) study is the only one to break the pattern by focusing on US restaurant firms. Furthermore, only two of the studies listed (Skogsvik, 1987; Lennox, 1999) were conducted outside the US.

An aspect of bankruptcy prediction research that is not covered in *Table 1* is the number of variables included in each model. In the compilation by Bellovary et al. (2007), a total of 752 different variables were identified in the bankruptcy prediction literature and the number of factors considered in a single

[11] For the purposes of this study, 'institutional context' refers to the economic, political and juridical setting a firm operates in as well as its industry association.

[12] For a more comprehensive compilation of bankruptcy prediction literature, see e.g. Bellovary et al. (2007).

Author(s)	Country	Industry focus	Statistical method	Hold-out sample	Error rates
Beaver (1966)	US	Unfocused	Univariate discriminant analysis	Yes	**13-24%** for the best predictive ratio 1-5 years before failure
Altman (1968)	US	Manufacturing firms	Multivariate discriminant analysis	Yes	**12.6%**, 1 year before failure
Ohlson (1980)	US	Unfocused	Logit analysis	No	**14.9%,** 1 year before failure
Zavgren (1985)	US	Manufacturing firms	Logit analysis	Yes	**31%** for all forecasting horizons, 1-5 years
Skogsvik (1987)	Sweden	Mining and manufacturing firms	Probit analysis	Yes	**16.0-28.8%** for current cost ratios 1-6 years before failure and **16.7-26.7%** for historical cost accounting ratios 1-6 years before failure
Lennox (1999)	UK	Unfocused	Probit analysis	Yes	**18.5%**, 1 year before failure
Kim & Gu (2006)	US	Restaurant firms	Logit analysis & MDA	Yes	**7%** for both models, 1 year before failure

Table 1 – Comparison of bankruptcy prediction studies

model ranged from one to 57 (Bellovary et al., 2007, p. 7). However, studies have been more consistent with regard to the average number of factors, which has been between 8-10 over the past 40 years (Ibid.). Bellovary et al. furthermore observed that an increased number of factors not necessarily entails a better model in terms of predictive accuracy, and that more parsimonious models often are preferable.

The studies listed in *Table 1* have all made use of accounting-based financial ratios to predict business failure. There exists a second school of prediction models that are based on market information instead of accounting information. Agarwal & Taffler (2008) acknowledged the recent growth in popularity of marked-based models and intended to compare the prediction performance of market-based and accounting-based bankruptcy prediction models in the UK. The idea with market-based models is that, given efficient markets, stock prices will reflect accounting information as well as all information *not* contained in financial statements (Agarwal & Taffler, 2008). Despite the theoretical appeal of market-based models, Agarwal and Taffler found that the predictive accuracy was similar to that of accounting-based models. However, the main reason for not examining market-based models in the present study is not related to the prediction accuracy but rather to the fact that only few European football clubs are listed on stock exchanges[13]. Gerritsen (2015) had a similar line of reasoning around the applicability of market-based models on football clubs. Accordingly, the financial ratios used in the model of the present study are entirely accounting-based.

It has been suggested that future research should focus on the application of existing models rather than the development of new ones (Bellovary et al., 2007). At the same time, *Subsection 2.1* hinted at the special conditions present in the football industry and the applicability of models developed for vastly different industries is questionable. *Subsection 2.3* elaborates on this by focusing on literature that has attempted to understand financial distress in the context of football.

2.3 Bankruptcy Prediction Models in Football Literature

As indicated in the previous subsection, accounting-based bankruptcy prediction models can be fundamentally different with regard to industry foci, statistical methods, countries of study, time periods and several other aspects. Applications of existing accounting-based bankruptcy prediction models on the European football industry have been made. For instance, Barajas and Rodríguez (2014)

[13] In the Big Five leagues, four clubs are publicly traded.

applied Altman's (2000) re-estimated Z-score model for private firms on Spanish first and second division clubs during the period 2007-2011. The aim of the study was to examine the financial situation of Spanish clubs and to determine the infusion of capital necessary to classify all clubs as 'surviving' according to the Z-score model (Barajas & Rodríguez, 2014). Although Barajas and Rodríguez acknowledged the limitations of using the Z-score model outside the US and for a different industry, they argued that it was useful in order to gain a rough understanding of the financial situation in Spanish football. The results indicated that Spanish clubs would have to issue equity in excess of €900 million for all clubs to be classified as surviving (Ibid.).

Whereas Barajas and Rodríguez (2014) used a bankruptcy prediction model for classification purposes, Gerritsen (2015) tested the accuracies of prominent prediction models when applied to football clubs. More specifically, Gerritsen compared the predictive accuracies of Ohlson's (1980), Zmijewski's (1984) and Altman's (2000) models on a sample of Dutch football clubs in the period 2010-2014. Gerritsen found that Zmijewski's model had the highest accuracy with an error rate of 34% for one-year forecasting periods. The corresponding error rates for Altman's (2000) re-estimated Z-score model and Ohlson's (1980) model were significantly higher at 57% and 81%, respectively. However, Gerritsen did not disclose the amount of misclassified distressed and non-distressed clubs. Consequently, it is uncertain if the results of his study are directly comparable to the accuracies listed in *Table 1*, where an arithmetic average of the two error rates are used. Gerritsen concluded, given the low accuracies, that none of the models were appropriate for application on the Dutch professional football industry. The reason for the models' low predictive accuracies was argued to be that professional football clubs differ from regular companies in that they are win-maximizing instead of profit-maximizing and the models were developed for profit-maximizing firms (Gerritsen, 2015). Furthermore, Gerritsen maintained that clubs took on excessive risks in order to maximize wins and that the only reason clubs tended to survive was that they were backed by governments or certain supporters, which is in line with the notion of benefactor owners presented in *Subsection 2.1*.

Barajas and Rodríguez (2010) and Szymanski (2012) developed models specifically for the football industry. However, these models were not designed to *predict* bankruptcies but rather to understand the reasons for why they occur. Barajas and Rodríguez (2010) studied the Spanish first and second division and attempted to find a variable or a set of variables that could explain why a club was in insolvency proceedings in 2008. Barajas and Rodríguez employed a logit regression to achieve this objective but underlined the small sample size as a limitation. The study used a single year of data and the sample consisted of only 35 clubs, six of which were in insolvency proceedings (Barajas and Rodríguez, 2010). When running the logit model with all variables considered, they found that the only significant variable was *Division*. However, when testing *Division* on its own it was insignificant. Hence, the conclusion arrived at by the authors was that none of the variables could explain why clubs become insolvent in Spain (Ibid.).

Szymanski's (2012) model did not suffer from the same limitations regarding time frame and sample size as Barajas and Rodríguez' (2010) model. Szymanski examined the causes of insolvency for English football clubs by studying data from the first four divisions in England during the period 1974-2010. In total, 67 clubs had entered insolvency proceedings during the chosen time frame (Szymanski, 2012). A set of probit regressions including two financial ratios and several non-financial ratios was run and the results showed that "negative shocks" (Ibid., p. 3) with regard to demand and productivity were the most significant reason for insolvencies in English football. These shocks could either be caused by an unexpected drop in revenues or by an underperformance of the club's players (Ibid.). The accuracies with which the model could predict that clubs become insolvent was not tested and in order to apply Szymanski's (2012) model, revenue and player wages information for *all* professional English clubs is required. Consequently, practical application of Szymanski's model is a lengthy process that requires significant amounts of information.

Scelles et al. (2016) developed a model similar to that of Szymanski (2012) and examined the reasons for insolvencies in the top three divisions of French football between 1970 and 2014. A difference between the models of Scelles et al. (2016) and Szymanski (2012) was that Scelles et al.'s model did not consider financial ratios. Scelles et al. found that a decline in match attendance, similar to the demand shocks discussed by Szymanski (2012), was a significant reason for why French clubs became insolvent. The model suffers from the same application difficulties as Szymanski's (2012) model.

2.4 Research Question

The European football industry constitutes a difficult field of application for bankruptcy prediction models. Previous researchers have analyzed the financial situation in various European countries and concluded that many clubs have unsustainable business models with excessive risk-taking, over-indebtedness and net losses due to lower-than-expected revenues and exorbitant player wages. Attempts of applying existing bankruptcy prediction models displayed low predictive accuracies

(Gerritsen, 2015) and the development of a model tailored to the football industry proved challenging (Barajas & Rodríguez, 2010). On the other hand, studies examining the reasons for financial distress have been more successful (Scelles et al., 2016; Szymanski, 2012), and the results have indicated that non-financial ratios are essential in order to understand financial distress in the context of football. Researchers have suggested the development of a model specifically designed for the football industry (Gerritsen, 2015) that uses accounting-based information to predict financial distress (Barajas & Rodríguez, 2010; Barajas & Rodríguez, 2014). Consequently, a gap in the literature is observed, and the present study aims to fill this gap by developing a prediction model for financial distress in European football clubs. The model in question is expected to consist of a mélange of accounting-based financial variables from the bankruptcy prediction literature and financial as well as non-financial variables identified as particularly important for football clubs. In developing the model, the research question of the present study is answered, namely: *Is it possible to develop a model that uses publicly available information to predict financial distress for European football clubs?*

Statistical analysis is employed to distinguish surviving clubs from distressed clubs and this distinction can become obscured by the special characteristics inherent in the football industry. As an example, the fact that several football clubs show consecutive years of negative equity and losses while at the same time being considered non-distressed (see *Subsection 5.1*) should obfuscate or at least understate the differences between surviving and distressed clubs. Thus, the predictive accuracy achieved by the model in this study is expected to be lower compared to the accuracies of previous bankruptcy prediction models developed for profit-maximizing companies in more traditional industries.

3. Methodology

This section provides the reader with a description of the statistical model used in this study. It begins with a brief discussion on the probit model, which is followed by an elaboration on the variable selection process. Subsequently, the model validation procedure is discussed in-depth, after which the section concludes by presenting a calibration formula for obtaining unbiased probabilities.

3.1 Designing a Model for the Football Industry

3.1.1 Probit Model

The statistical model of choice in this study is a probit regression model. Given that the model is estimated on an unbiased sample, the probabilities of financial distress generated by logit/probit models can be used in real-world decision contexts (Skogsvik & Skogsvik, 2013). As such, these types of models are well-suited and frequently used for predicting financial distress (e.g. Kim & Gu, 2010; Lennox, 1999; Ohlson, 1980; Skogsvik, 1990; Zavgren, 1985; Zmijewski, 1984). As discussed in *Subsection 2.2*, other statistical methods such as multivariate discriminant analysis are based on assumptions that rarely are fulfilled (Lennox, 1999; Ohlson, 1980; Skogsvik, 1990; Zavgren, 1985). Furthermore, models that only generate dichotomous classifications, i.e. distressed or surviving, and not likelihoods of failure, are less useful in practical application settings (Zavgren, 1985).

The underlying distributional assumption on which the probit model is based, the cumulative normal distribution, is similar to that of the logit regression except that the tails of the distribution are slightly thinner and the center is slightly denser (Dey & Astin, 1993). Given that the cut-off value for classification as failed or surviving (discussed further in *Subsection 3.2.2*) is expected to be in the vicinity of the middle of the distribution, the probit model is chosen for this study. However, inferences drawn from applications of logit and probit models, ceteris paribus, should be closely consistent (Lennox, 1999). The resulting index value of the probit regression (assigned with G in *Equation 1*) is converted to a probability of distress using a normal distribution table, which ensures that the probability derived from an application of *Equation 1* is between 0 and 1 (Wooldridge, 2012). The probit function is:

$$\mathrm{P}(\mathrm{Y} = 1 | X_1, X_2, \ldots, X_k) = \mathrm{G}(\beta_0 + \beta_1 X_1 + \ldots + \beta_k X_k) \qquad (1)$$

where P = Probability

Y = The dependent variable

X_i = The value of an independent variable

G = The standard normal cumulative distribution function

β_i = Coefficient of an independent variable

3.1.2 Variables

The first step in the development of the probit model for predicting financial distress is to determine which independent variables to include. In estimating the model on the sample of clubs, these predictor variables are given coefficients that, when applied to a club, are intended to predict the probability of distress. In the present study, three different groups of independent variables are identified. The first group consists of financial variables that have been prominent in previous bankruptcy prediction studies. The second group is financial and non-financial variables that have been identified as relevant in the football literature. The third and last group consists of non-financial variables related to on-field performance and other characteristics of football clubs that were developed by the authors of this study.

In the field of accounting-based bankruptcy prediction, ratios that capture various aspects of the financial condition of companies are used as independent variables. These ratios can be divided into different categories. Skogsvik (1987) identified "profitability, cost structure, capital turnover, liquidity, asset structure, financial structure, [and] growth" (p. 344) as the broader categories that the financial ratios used in his study were grouped into. Prominent bankruptcy prediction models have generally made use of similar sets of financial ratios (see Bellovary et al., 2007, p. 7 & p. 42). All the financial ratios considered in the present study can be classified under one of the categories identified by Skogsvik (1987). The financial ratios from bankruptcy prediction studies were derived from the models of Altman (1968, 2000), Beaver (1966), Ohlson (1980), Skogsvik (1987), and Zavgren (1985).

Subsection 2.1 highlights the ways in which football clubs are different compared to other types of companies. Clubs are typically characterized by a distinct revenue and cost structure as well as a high proportion of intangible assets in the form of players on the balance sheet. Consequently, some of the ratios derived from previous bankruptcy prediction literature, which has often been focusing on manufacturing and other industrial companies, are ill-suited for football clubs. As an example, football clubs generally hold insignificant amounts of inventory. Hence, ratios such as inventory turnover are expected to be less relevant for clubs than for other types of companies where inventory is more prominent. To identify a set of financial ratios particularly relevant to the football industry, the studies of Barajas and Rodríguez (2010), Beech et al. (2010), Marcinkowska (2013), Mourao (2012), and Szymanski (2012) are examined. An example of a ratio that was identified as relevant to football clubs is *Wages/Revenues*, a ratio that gives an indication of the percentage of revenues that is spent on football clubs' major cost item: player wages. Previous studies (Barajas & Rodríguez, 2010; Szymanski, 2012) acknowledged the relevance of this ratio as an indicator of the financial health of clubs.

Given that the primary objective of clubs is to maximize wins on the field as discussed in *Subsection 2.1* and that they do not operate with the same logic as regular, profit-maximizing firms, the authors of the present study expect financial indicators to be less significant for football clubs. The fact that clubs can exhibit several consecutive years of losses but still be able to perform at the top of their divisions highlights this notion (Kuper & Szymanski, 2014). Consequently, in addition to the financial indicators identified thus far, non-financial indicators that are related to on-field performance and club characteristics are examined as well. These ratios are based on the studies of Buraimo et al. (2005), Barajas and Rodríguez (2010), Beech et al. (2010), and Szymanski (2012). Such indicators include, e.g., the division a club is playing in and dummy variables for being relegated or promoted.

In addition to the variables identified as relevant due to their prevalence in previous football or bankruptcy prediction literature, two types of variables are the construct of the authors of this study. The first of these types is related to whether or not the club has at least one 'star player', as defined by the best 120 players in a ranking provided by EA Sports' FIFA video game (FIFA Index, 2017). The reason for including a variable for star players is that clubs with such players in expectation should have a higher degree of on-field success. Furthermore, clubs with star players should have higher revenues, partly due to the on-field success and partly due to the popularity of the player, giving rise to increased ticket and merchandise sales. Despite the observation that the premiums paid for the best players exceed the additional revenue they entail (Boeri & Severgnini, 2012), the authors of this study expect the star player variable to indicate that clubs with such players have a lower probability of distress. The second type of variable is a dummy variable for Spanish clubs. The reason behind the inclusion of this variable is that, per Spanish law, the national football association in Spain cannot penalize clubs, by deducting points or relegation, for entering insolvency proceedings as this would inhibit future revenue potential of the distressed club (Barajas & Rodríguez, 2014). This feature, that to the best knowledge of the authors is unique to Spain, makes the entry into insolvency proceedings less detrimental to Spanish clubs and constitutes the reason for the Spain dummy variable.

The variable identification and development process led to 31 variables that were initially identified as relevant and applicable in the model based on previous research. In addition, five variables were the construct of the authors of this study. The variables tested in the model and their expected effect on distress are presented in *Appendix 1*.

3.1.3 Arriving at a Parsimonious Model

In order to reduce the number of predictors and arrive at a more parsimonious model consisting of a smaller set of uncorrelated variables, principal component analysis ('PCA') is employed. PCA is a variable reduction procedure that aids in reducing variables that are highly correlated and that essentially are measuring the same construct by merging several variables into components (O'Rourke, Hatcher & Stepanski, 2005). Given that PCA is designed for continuous, non-categorical variables (Niitsuma & Okada, 2005), dummy variables are excluded from this process. PCA creates components by optimally weighting and combining different variables to capture as much of the variance in the underlying data set as possible (O'Rourke et al., 2005).

The *.pca* command in the statistical software Stata results in orthogonal, or uncorrelated, components and is the principal component analysis method of choice in this study (Stata, 2013). The components explain varying degrees of the total variance in the underlying data set. Consequently, some components are accounting for an insignificant amount of the total variance and attempts to limit the number of components are made. This is done by applying the eigenvalue-one criterion, which makes use of the eigenvalues given to the components in the PCA, where an eigenvalue is a representation of how much of the total variation is explained by the component (O'Rourke et al., 2005). The rationale behind the eigenvalue-one criterion is that each of the considered variables contributes with one unit, or eigenvalue, of variance to the total variance of the data set. Thus, a component with an eigenvalue larger than one captures a greater amount of variance than any one variable on its own (Ibid.). Conversely, components with eigenvalues below one capture less of the total variance than any one variable and are not deemed significant enough to be retained (Ibid.).

To simplify the interpretation of components, a statistical technique called 'rotating' is used. By rotating the components, the coefficients of the variables in each component approach zero or one, leading the variables in a component to become more distinct and consequently easier to interpret (Abdi & Williams, 2010). There are two distinct categories of rotations, orthogonal rotations and oblique rotations. Within each category, several different rotation methods exist (Stata, 2013). Whereas orthogonal rotations, with the varimax rotation being the most prominent method, restrict the components to be uncorrelated, oblique rotations, with the promax rotation being the most prominent method, allow for some correlation (Abdi & Williams, 2010). For the purposes of this study, several rotations of both orthogonal and oblique nature are performed and the one that provides the simplest structure for interpretation is chosen. The approach of selecting which ratios to include in the final model from the components obtained in the principal component analysis is done by applying a similar variable selection process to that of Skogsvik (1987). This approach involves choosing the variable that has the highest correlation with each component to be included in the model, while simultaneously controlling that the correlations (in absolute values) between the selected variables are below 0.5 (Skogsvik, 1987). The reason for reducing the number of variables is that the model becomes more parsimonious and easier to apply in practice.

Once each of the relevant components has been assigned a variable with which it is most highly correlated, the probit regression is run including the variables determined through the PCA as well as the dummy variables determined to be relevant in the variable selection process. As some of the included variables are statistically insignificant, there is room to remove overabundant variables to make the model more parsimonious. In order to select which variables to include in the final model, a backward elimination procedure is used. In backward elimination, the starting point is to include all relevant independent variables in the model, after which the least significant variable is dropped until only variables with p-values below a predetermined cut-off value are included (Faraway, 2014). Given that the objective of the present study is to arrive at a parsimonious model that simultaneously explains a sufficiently large amount of the variation in the data, the cut-off value is chosen with this trade-off in mind. When the principal component analysis shows that two variables have almost equally high correlations to a certain component and the variable with the highest correlation is being dropped in the backward elimination process due to having a p-value above the chosen cut-off value, the variable that had the second highest correlation with the component is tested. Depending on the p-value of the newly included variable, it is either dropped or kept in the model. The model derived from the completion of the backward elimination process is the final model presented in *Section 5*.

3.2 Assessing the Model's Quality

3.2.1 Application Samples

Ideally, the model should be validated by applying it to a completely new set of observations that was not used for the estimation of the model, i.e. a true hold-out sample (Bellovary et al., 2007). However, this procedure is seldom feasible when sample sizes are small and many previous bankruptcy prediction studies have instead applied the 'jackknife' procedure or alternative versions thereof (Ibid.). The rationale behind the jackknife procedure is that single observations or groups of observations are held out of the estimation of the model and subsequently predicted as surviving or distressed (Bellovary et al., 2007; Skogsvik, 1987). This entails that there is no overlap between the observations used for the estimation of the model and the observation on which the prediction accuracy is tested.

A similar validation approach to that of Skogsvik (1987) is applied in this study, which involves two steps. First, the sample of analysis is used for model estimation, after which the entire sample is also used for model validation where the target is to classify clubs as surviving or distressed. Thus, at this stage there is no hold-out sample. In terms of validity, this analysis of the estimation sample is questionable (Eisenbeis, 1977; Joy & Tollefson, 1975; both cited by Skogsvik, 1987). Yet, several previous studies present only these kinds of results and do not take hold-out samples into consideration (Skogsvik, 1987; Bellovary et al., 2007). Therefore, this step will enable the results of this study to be compared to studies that only used this approach.

Second, the jackknife procedure is applied, which allows for cross-sectional validation (Skogsvik, 1987). The jackknife procedure involves dividing the sample of analysis into separate groups. In the present study, the total sample is divided into two equally sized groups. Distressed and surviving clubs are drawn proportionately to the ratio of distressed and surviving clubs in the total sample to form the two subsample groups. The model is re-estimated so that the estimated coefficients are based on one of the subsample groups. The re-estimated model is validated by applying it to the second group that was held out from the estimation of the model and that thus serves as a hold-out sample. The process is thence repeated but reversely so that the group that was held out from the estimation now serves to estimate the model and vice versa.

3.2.2 Classification

In terms of predictive accuracy, there are two types of errors associated with predicting financial distress. The first, 'type I' error, occurs when the model misclassifies a financially distressed firm as surviving and the second, 'type II' error, is the misclassification of a surviving firm as financially distressed. *Table 2* illustrates the error matrix.

		Classification	
		Distress	**Survival**
Situation	**Distress**	Correct	Error Type I
	Survival	Error Type II	Correct

Table 2 – Error matrix

As indicated earlier, the model is estimated and each club is given a probability of failure that depends on the coefficients of the model as well as on the club's individual scores on each of the variables. The next step is, given the probabilities provided by the model, to classify the clubs as surviving or distressed. To achieve this, a cut-off value needs to be chosen, above which clubs are classified as distressed and below which clubs are classified as surviving. By setting a low cut-off value, fewer type I errors would occur and, conversely, by setting a high cut-off value, fewer type II errors would occur. Ohlson (1980) mentioned that previous research in the field of bankruptcy prediction found that the best model was the one that minimized the total errors, which is consistent with the 'empirical approach', as discussed by Skogsvik & Skogsvik (2013). Choosing a cut-off value according to the empirical approach entails minimizing the average error rate or error cost (Skogsvik & Skogsvik, 2013). The approach chosen in this study is to determine the optimal cut-off point based on minimizing the arithmetic average error rate ($rate(\bar{e})$, given by *Equation 2*), or the weighted-average error rate ($rate(\bar{e})'$, given by *Equation 3*). These error rates are (Skogsvik & Skogsvik, 2013, p. 33):

$$rate(\bar{e}) = [rate(e_I) + rate(e_{II})]/2 \quad (2)$$

$$rate(\bar{e})' = prop * rate(e_I) + (1 - prop) * rate(e_{II}) \quad (3)$$

where $rate(e_I)$ = error rate type I

$rate(e_{II})$ = error rate type II

$prop$ = the proportion of distressed clubs in the estimation sample

Using this approach, no distinction is made between the costs associated with type I and type II errors. It is worth noting that type I errors generally are costlier than type II errors (Bellovary et al., 2007). However, following the line of argument of Zavgren (1985), both error types are perceived as important. Since previous literature is ambiguous regarding the true costs of the two types of errors, the focus in the present study is to find a model that minimizes weighted average error rates as given by $rate(\bar{e})'$ instead of considering costs of misclassification. However, when comparing to previous research, $rate(\bar{e})$ is of relevance. Many of the earlier bankruptcy prediction studies used a one-to-one matched-pairs sampling approach which entailed 50% distressed firms and 50% surviving firms in the sample. For that reason, researchers have often turned to this arithmetic average error rate in order to compare to other studies.

To answer the research question if an accounting-based prediction model can be beneficial for predicting financial distress for football clubs, the results of the accuracy testing are also compared to naïve decision rules. Skogsvik (1987, p. 52) proposed two such decision rules:

Rule 1: Classifying all clubs according to the classification that a priori has the highest probability of being correct.

Rule 2: Classifying clubs randomly, but in proportion to the a priori probabilities of both outcomes, as financially distressed or surviving.

Given the a priori probability of distress in the sample of clubs in this study (discussed in detail in *Subsection 4.2*), application of Rule 1 would entail classifying all clubs as surviving. The real-world probability of distress is even lower, suggested to be in the vicinity of 2-3% for English and French football clubs (Szymanski 2012; Scelles et al. 2016). Rule 1 would, consequently, always classify all clubs as surviving. Hence, such a rule does not attempt to predict financial distress and the practical usefulness is limited. As such, it receives no further attention in this study. The second naïve decision rule classifies clubs randomly, but proportionally according to the a priori probabilities. As an example, by supposing that half of the clubs are distressed and half of the clubs are surviving in the sample, the second naïve decision rule would randomly classify half of the clubs as surviving and the other half of the clubs as distressed. Given that Rule 2 predicts clubs as both distressed and surviving, based on the a priori probabilities, it serves as a relevant benchmark for the prediction model in this study.

A final way to contextualize the accuracy achieved by the model developed in the present study is to apply an existing prediction model to the same sample of clubs. Given that Gerritsen (2015) utilized several prominent prediction models on a sample of Dutch clubs and found that Zmijewski's (1984) model had the highest prediction accuracy, Zmijewski's model is also applied to the sample used in this study. Worth acknowledging is that the main reason for the development of Zmijewski's (1984) model was to examine the existence of biases related to non-random samples and not for predictions, which is why the model was not listed in *Subsection 2.2*. Zmijewski's model that was applied in Gerritsen's (2015, p. 19) study is:

$$Zmijewski = -4.3 - 4.5X_1 + 5.7X_2 + 0.04X_3 \qquad (4)$$

where $Zmijewski$ = the dependent variable, for values < 0.5 clubs are classified as surviving and for values ≥ 0.5 clubs are classified as distressed

X_1 = Net income/Total assets

X_2 = Total liabilities/Total assets

X_3 = Current assets/Current liabilities

The application of the model given by *Equation 4* on the sample of this study is presented in *Subsection 6.1*. In expectation, Zmijewski's model should have a lower prediction accuracy than the model developed in this study given that it was developed for a completely different set of firms over 30 years ago.

3.3 Real-world Application

When a bankruptcy prediction model is based on a non-random sample that does not represent the real-world proportion of distressed and survival firms, as is the case in the present study, the probabilities provided by the model are biased (Skogsvik & Skogsvik, 2013). Even though the probabilities are biased, the ranking of firms obtained from such a model is still correct (Ibid.). However, in real-world decision contexts, unbiased probabilities are required. Unbiased probabilities can be achieved by re-estimating the model on a sample consisting of a real-world proportion of distressed and surviving firms, a process that is cumbersome and time-consuming. A second option is to utilize the adjustment formula provided by Skogsvik & Skogsvik (2013), which is used to calibrate the sample-based probabilities. The calibration formula (Skogsvik & Skogsvik, 2013, p. 32) is given by:

$$P_{fail}^{(adj)} = \left[1 + \left(\frac{1-\pi}{\pi}\right) * \left(\frac{prop}{1-prop}\right) * \left(\frac{1-P_{fail}^{(prop)}}{P_{fail}^{(prop)}}\right)\right]^{-1} \qquad (5)$$

where π = a priori probability of failure in the population

$prop$ = the proportion of distressed clubs in the estimation sample

$P_{fail}^{(adj)}$ = estimated unbiased probability of failure, calibrated for the probability of failure in the population

For the formula to be applicable, it is required that "the samples of bankrupt and survival firms constitute random drawings from the sub-populations of bankrupt and survival firms, respectively" (Skogsvik & Skogsvik, 2013, p. 32). Caution is advised regarding the application of the formula on the results of the model in this study. The data sample used in the present study does not fulfill the conditions necessary since the data for distressed clubs was selected on the basis of availability, as described in the following section.

4. Data

This section commences by delineating the definition of financial distress that is applied in this study, after which a discussion of the data sample follows.

4.1 Definition of Financial Distress

To clarify what the model developed in this study is predicting, the definition of the notion 'financial distress' used to classify the clubs in the sample must be outlined. In previous literature, there is a broad scope of definitions used, ranging from actual bankruptcy filings to the inability to pay a preferred stock dividend (Karels & Prakash, 1987). The criteria used to identify financially distressed clubs in this study are based on the first and third criterion used by Skogsvik's (1990, p. 142):

1. Bankruptcy and/or a composition agreement.
3. Receipt of a substantial subsidy provided by the state [without which financial distress would ensue].

With creditors that are reluctant to let clubs go bankrupt and non-profit seeking owners, European football clubs are in a special situation with regard to financial distress. Clubs, just like regular companies, enter into insolvency proceedings (Barajas & Rodríguez, 2014; Beech et al., 2010; Szymanski, 2012), e.g. administration in the UK or the Ley Concursal in Spain. These processes are analogous to Chapter 11 bankruptcy in the US (Barajas & Rodríguez, 2010). Entering such a process corresponds to Skogsvik's first criterion, which is also applied in this study. The sample of clubs in this study stretches over multiple European legislations and the judicial systems differ between countries. Taking a broad perspective, the triggers of insolvency and the insolvency proceedings are similar in all of the countries examined in this study (The Law Firm Network, 2013). Generally, a firm is obligated to enter insolvency proceedings if it fails to fulfill either the cash-flow test, referred to as actual insolvency because the firm finds itself in a liquidity crisis, or the balance sheet test, referred to as technical insolvency because total liabilities exceed total assets (Margret, 2002). Both the directors of the firm as well as the creditors can file for insolvency proceedings should the company be insolvent (The Law Firm Network, 2013). However, the football industry presents some irregularities in this regard. Taking Chelsea F.C.[14] as an example, the club's liabilities have exceeded its assets since at least 2006 (Bureau van Dijk, 2017). Hence, the club is failing the balance sheet test. However, since all parties trust the continuous supply of funds by the club's benefactor to ensure liquidity, neither the creditors nor the directors intend to file for insolvency proceedings (Beech et al., 2010). Similarly, reluctance from owners and directors to make clubs enter insolvency proceedings means that the cash-flow test is not always binding either. Several clubs are unable to pay their taxes owed to the government or to pay their players (Barajas & Rodríguez, 2010; Morrow, 2014). As an example, Southend United[15] were able to avoid insolvency proceedings despite being unable to pay taxes (see

[14] Chelsea Football Club Limited.
[15] Southend United Football Club Limited(The).

Appendix 2 for all distressed clubs included in the study). However, failing to pay players or taxes and consequently being in the state of actual insolvency, *is* regarded as distress in this study.

Receiving continuous financial contributions from a benefactor, as is the case for Chelsea F.C., is *not* seen as being financially distressed in the present study, despite the fact that the club has negative equity. However, as presented in criterion three of Skogsvik (1990), receipt of a one-time financial subsidy by the state, here extended to a company or private persons, for a club that would have faced immediate bankruptcy procedures otherwise, is determined as a criterion for financial distress. Feyenoord Rotterdam[16] from the Netherlands presents a case in point when it was bailed out by a group of investors in return for a 49% stake in the club after accumulating excessive debts and being threatened with insolvency proceedings in October 2010 (*Appendix 2* lists all clubs that have encountered such problems).

Finally, a club that is penalized by its national football association or the UEFA for insufficient financial performance is also classified as distressed. Malaga C.F.[17] is an example of a club penalized by the UEFA for breaking UEFA's financial fair play rules in 2012 and Nottingham Forest[18] broke English financial fair play rules in 2014. By breaking these rules, clubs receive financial penalties, transfer embargos, point deductions, or similar penalizations (see *Appendix 2*) that affect future revenue potential.

To summarize, a club is considered to be distressed if it enters insolvency proceedings, fails to pay any of its creditors due to a lack of liquid assets, is bailed out by one or several benefactors before insolvency proceedings can be initiated, or is penalized by a national or European football association for having a poor financial situation.

4.2 Data Sample

The sample of clubs in the study is designed to represent the population of the major European football leagues in the period 2006-2016. The UEFA ranking for European club competitions in the 2016/17 season (UEFA, 2017) provides the basis for the data sample selection. The leagues commonly referred

[16] Feyenoord Rotterdam N.V.
[17] Malaga Club de Futbol, Sociedad Anonima Deportiva.
[18] Nottingham Forest Football Club Limited.

to as the 'Big Five', namely England, Spain, Germany, France, and Italy, are at the forefront of the ranking in terms of on-field performance, but also in terms of revenues (Hembert et al., 2010). The ranking is followed down to Scotland on rank 23 and all the leagues in between are considered. In addition to the Big Five, the leagues in the following countries had at least some financial information available: Austria, Belgium, Czech Republic, Denmark, Netherlands, Poland, Portugal, Scotland, Sweden and Switzerland. In other countries, namely Belarus, Croatia, Greece, Romania, Russia, Turkey and Ukraine there was no financial information available and they were hence excluded from the study. Similarly, Israel was not included because, despite being part of the UEFA, it economically belongs to Asia, a region that is not covered by the database that is used in this study.

Apart from Switzerland, all the included countries are members of the European Union. Consequently, the economic, legal and political environments in the studied countries can be expected to be relatively consistent. Switzerland is assumed to be sufficiently similar to the other countries included in the sample in these regards. The authors of this study acknowledge that these assumptions are not unquestionable, but maintain that assumptions of this kind must be made in order to have a sufficiently large sample of football clubs. The limitations of making such assumptions are addressed in *Section 7*.

Furthermore, the operations of football clubs are expected to be similar in all examined countries, with most of the revenues stemming from sales of broadcasting rights and with player wages constituting the bulk of the expenses (Garcia-del-Barrio & Szymanski, 2009). Given that the national football associations are members of the UEFA, the competitive format is similar in the studied countries with relegation and promotion procedures for the worst and best teams of a division. The clubs in the first division in the selected countries are examined in each season between 2006 and 2016. However, due to the relatively larger commercial sizes of the leagues in the Big Five countries, the first *and* second divisions are considered for those countries. The data sample in this study thus includes all clubs that played at least one season in any of the aforementioned leagues during the studied time period for which data is available. Given that the clubs examined in this study are not randomly selected, statistical inferences cannot be made about the population of all European football clubs but only about the sample (Smith, 1983).

The financial data was primarily obtained through the Amadeus database from Bureau van Dijk which provides financial information for millions of public and private European companies in a standardized format that allows for comparisons between countries (Bureau van Dijk, 2017). It is important to acknowledge, however, that the standardized format could entail less detailed financial information since line items are amalgamated and transformed to fit Amadeus' standardized format. The database provides general company information as well as income statement and balance sheet of companies included in the database. No statements of cash flow are available. The maximum number of years that Amadeus provides financial statements for any given company is ten and, consequently, the studied time period was influenced by this limitation. For various reasons, some clubs were not included in the Amadeus database or parts of the information was missing. Where possible, attempts to retrieve the information for clubs excluded from Amadeus were made through other channels. This includes data transcribed manually from Companies House, which is the registrar of companies in the United Kingdom (Companies House, 2017), from Bundesanzeiger, which is the registrar of companies in Germany (Bundesanzeiger, 2017), or, if available, from the financial statements directly retrieved from a club's webpage. The legal forms represented in the sample are mostly public limited companies and private limited companies[19] (Bureau van Dijk, 2017). Most private and all public limited companies are obliged to publish financial statements. However, a large number of clubs also operate as associations, especially in Germany, and do not have to publish financial statements. This data unavailability is expected to affect the sample and consequently the clubs for which inferences can be made.

In addition to financial data, non-financial information related to the on-field performance of the clubs is used. The on-field performance-related data such as division in a respective year and league promotion or relegation was collected from *transfermarkt.com* (2017). Furthermore, Eurostat's Regional Yearbook (2015) was used to retrieve information on the population of the clubs' municipalities. Finally, a ranking of the world's best football players was retrieved from *fifaindex.com* (2017), an index that lists the players of EA Sports' FIFA game by their total score for all versions of the game from 2005 onwards. The FIFA game player ranking is based on ratings from a network of 9,000 reviewers (Lindberg, 2016).

Having identified all clubs that make up the sample population, each club was controlled against the criteria for financial distress as outlined above. This was done through news research and through the Amadeus database that states if companies have been dissolved. Through this process, the month of

[19] Out of the final sample of 208 clubs, 121 clubs are public limited companies and 78 clubs are private limited companies. In addition, there are four clubs operating as non-profit organizations and five partnerships included in the sample.

financial distress was identified for all distressed clubs. The full list of distressed clubs is presented in *Appendix 2*. The surviving clubs included in the sample represent all clubs for which there is data available for at least three consecutive years during the period of analysis.

The sample of distressed clubs is not randomly selected. The 52 financially distressed clubs make up the entire population of clubs that became financially distressed during the period 2006-2016 for which there is at least one year of data before distress available. As suggested by the literature in *Subsection 2.2*, a matched-pairs sample was used in this study where the financially distressed clubs are matched against surviving clubs based on accounting period end and asset size. The accounting data used for surviving clubs was gathered for the same year as for the matched, distressed clubs. Despite the fact that Szymanski (2010) did not find the global financial crisis to have a significant impact on football clubs, studying distressed and surviving clubs in the same time period should mitigate the unobservable effects of the overall economic environment. Matching based on size should likewise control for inexplicit factors (Zavgren, 1985). Due to the access to a higher number of surviving firms, each distressed club is paired with three surviving clubs. Consequently, the total sample consists of 52 distressed and 156 surviving clubs.

When using a prediction model to predict financial distress, the input information for the model need to be available to the user. The example in *Figure 2* illustrates this.

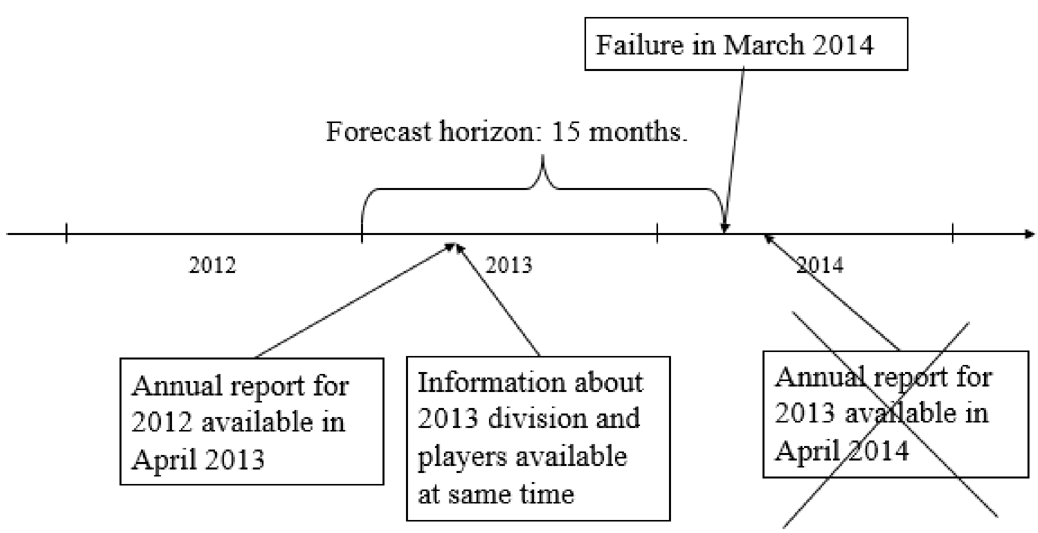


Figure 2 – The forecast horizon problem. Inspired by Ohlson (1980)

As *Figure 2* shows, the financial statements of 2013 are not yet available at the time of distress in March 2014. Hence, the financial statements of 2012 are the ones used to predict distress in March 2014. If this timing issue is not considered, there is a risk of *'back-casting'* (Ohlson, 1980, p. 110), where the prediction is based on information that became available after the club has already failed. To ensure consistency and avoid the risk of back-casting, the prediction model development in this study uses financial information of the period ending *at least* twelve months before distress, even if more recent information is available. The average length between the end of the reporting period of the financial statements used and the date of financial distress in this study is 18.8 months. This is relatively longer compared to other studies, examples being Skogsvik's (1987) average forecasting horizon of 13.2 months or Lennox' (1999) average forecasting horizon of 14 months. In expectation, a longer forecasting period should entail a lower predictive accuracy given that the information used and the distress event are further apart in time.

There is a risk that outliers might strongly influence the model (Wooldridge, 2012). At the same time, outliers provide information about the sample given that the underlying measurement is correct. Consequently, outliers were only excluded when an error in the data could be identified. The authors observed that errors in the data were present in four cases, resulting in the exclusion of four clubs from the sample. Those consisted of one distressed and three survival clubs. The final number of clubs presented above, 52 distressed and 156 surviving, is after outlier-exclusion.

5. Empirics and Analysis

This section begins with presenting the descriptive statistics of the sample. Thereafter, the final model is shown and the included variables are examined. An analysis of the accuracy of the model follows and it is compared to previous bankruptcy prediction studies. The section concludes with the same procedure being repeated for the two-year model.

5.1 Descriptive Statistics

Out of the 52 financially distressed clubs for which data is available, Spanish clubs constitute roughly 27% (*Table 3*). The number of identified clubs is strongly affected by data availability. Hence, inferences about a generally higher likelihood of distress in one country compared to another cannot be made solely from considering the number of distressed clubs in each country.

	Czech Republic	England	France	Germany	Italy	Netherlands	Poland	Portugal	Scotland	Spain	Total
distressed	1	7	2	3	11	3	4	3	4	14	52

Table 3 – The sample number of distressed clubs from each country

As *Table 3* shows, the number of distressed clubs in the sample is highest in Spain, followed by Italy and England. Given that the first *and* second divisions are examined in these countries, a larger number of distressed clubs can be expected. On the other hand, the second divisions in France and Germany are also included, yet the number of distressed clubs is significantly lower at two and three, respectively. No distressed clubs with available data are identified for the leagues in Austria, Belgium, Denmark, Sweden and Switzerland.

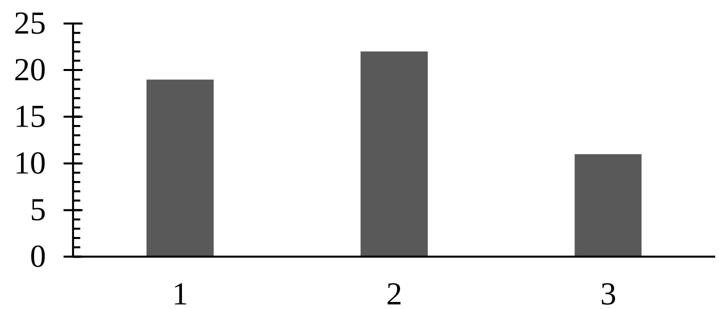


Figure 3 – Number of distresses observed per division

Contrary to what previous literature identified (Szymanski, 2012), there is a number of clubs that became distressed while playing in the top division of the respective countries (*Figure 3*). This is linked to the stricter definition of distress applied in this study. 19 clubs were playing in the first division at the time of distress and 22 clubs were playing in the second division at the time of distress. The clubs experiencing distress in the first division mainly originate from Spain (5), Portugal (3), Poland (3), and the Netherlands (3). While clubs in the first division of the Netherlands and Portugal experienced liquidity issues or received financial support to avoid entering insolvency proceedings, most of the clubs in Spain and Poland entered insolvency proceedings while playing in the first division (*Appendix*

2). Given that the sample includes clubs that played in any of the examined divisions during at least one season in the period 2006-2016, there are eleven clubs that became financially distressed while playing in the third division.

Looking at the descriptive statistics of the variables identified in *Subsection 3.1.2* (*Table 4*), there are several differences between distressed and survival clubs. The means for the variables of distressed clubs in the sample are generally 'worse' than for the surviving clubs in the sample, i.e. they lean towards the extreme that is expected (see *Appendix 1*) to indicate distress, with some exceptions for the division and movement variables. This is also observed when examining the medians of the ratios. Looking at specific variables, several of the characteristics of football club identified in *Subsection 2.1* can be recognized in the descriptive statistics of the sample. As an example, the ratios *Net Income/Total Assets* as well as *EBIT/Total Assets* are negative on average for surviving clubs at -0.08 and -0.16, respectively. As can be expected, the corresponding averages for the distressed clubs are significantly worse around -0.27 and -0.29, respectively. The medians are also negative for both surviving and distressed clubs. In addition, employee costs, which includes both player wages and administrative employee costs, represent the major cost item with 70% of revenues going to employees for surviving clubs and 82% of revenues going to employees for distressed clubs. For both groups this is beyond the 67%-mark that Késenne (2009, cited by Barajas & Rodríguez, 2010) established for a club to be identified as profit-maximizing. Thus, given our data sample, these statistics support prior researchers that argued that football clubs are not profit-maximizing entities. Furthermore, total liabilities are on average 59% higher than total assets in distressed clubs, which sheds light on the detrimental state that these clubs are in one year before the distress event. However, on average, total liabilities exceed total assets for surviving clubs as well, although at a lower rate of 10%. By looking at the median of the ratio, it is observed that liabilities exceed assets for the median of distressed club but not for the median of surviving club. One reason for the difference between mean and median for distressed clubs is OS Belenenses'[20] extreme ratio of 9.1.

[20] OS Belenenses - Sociedade Desportiva De Futebol, S.A.D.

	Mean		Median		Standard Deviation	
Variable	**Survival**	**Distress**	**Survival**	**Distressed**	**Survival**	**Distress**
Spain	0.1154	0.2692	0.0000	0.0000	0.3205	0.4479
Population (01.01.2012)	878,031	351,844	217,635	247,450	1,894,558	361,593
Division Y	1.5962	1.6731	1.0000	2.0000	0.7430	0.6484
Division Y-1	1.6346	1.6346	1.0000	1.0000	0.7795	0.6271
Division Y-2	1.6923	1.5385	2.0000	1.5000	0.8697	0.5760
Relegation Y-1	0.1026	0.1538	0.0000	0.0000	0.3044	0.3643
Promotion Y-1	0.1410	0.1154	0.0000	0.0000	0.3492	0.3226
Short-term Movement (2 years)	0.4808	0.4038	0.0000	0.0000	0.6765	0.5691
Short-term Movement (3 years)	0.6538	0.5769	1.0000	0.0000	0.8919	0.7758
Long-term Movement (5 years)	1.1154	1.1154	2.0000	1.0000	1.1414	0.9833
Star Player Y to Y-2	0.1603	0.0385	0.0000	0.0000	0.3680	0.1942
Star Player Y	0.1026	0.0192	0.0000	0.0000	0.3044	0.1387
Star Player Y-1	0.1218	0.0000	0.0000	0.0000	0.3281	0.0000
Star Player Y-2	0.1218	0.0192	0.0000	0.0000	0.3281	0.1387
BV Equity / BV Liabilities	0.4273	-0.1327	0.0753	-0.0448	1.5252	0.3688
Wages / Revenues	0.6699	0.8186	0.6883	0.7407	0.2654	0.4043
Owner's Equity / Assets	-0.0972	-0.5889	0.0700	-0.0477	0.9114	1.4248
Total Liabilities / Total Assets	1.0972	1.5884	0.9300	1.0477	0.9114	1.4251
Wages / Total Costs	0.5568	0.5956	0.5587	0.5532	0.1731	0.2275
Amortization / Total Costs	0.1170	0.1467	0.1214	0.0960	0.0836	0.1453
Long-term Debt / Revenues	0.3883	0.4386	0.0041	0.1098	0.9096	0.6916
Revenues / Total Liabilities	1.5738	0.7854	0.7922	0.4828	1.4421	0.8193
Total Costs / Revenues	1.2148	1.4070	1.1162	1.2946	0.3766	0.5861
Working capital / Total assets	0.0472	0.0607	0.0303	0.0553	0.2440	0.2007
Current Liabilities / Current Assets	2.0343	5.2008	1.5914	2.0792	2.4964	10.5357
Current Assets / Current Liabilities	1.4888	0.4977	0.6284	0.4820	7.0156	0.3249
Cash assets / Current Liabilities	0.2363	0.0628	0.0709	0.0168	0.3415	0.1307
Net Income / Total Assets	-0.0796	-0.2739	-0.0531	-0.1222	0.2842	0.4555
EBIT / Total Assets	-0.1605	-0.2928	-0.1228	-0.1723	0.3487	0.4304
Revenues / Total Assets	1.3388	1.0343	1.0267	0.5906	1.4724	1.0380
Change in Net Income	-0.0191	-0.1428	-0.0450	-0.1387	0.7184	0.7041
INTWO	0.3782	0.4808	0.0000	0.0000	0.4865	0.5045
OENEG	0.2949	0.5192	0.0000	1.0000	0.4575	0.5045
Equity / Fixed Assets	0.1032	-1.8183	0.1437	-0.1320	3.6345	8.0414
Change in Revenues	0.1798	0.0435	0.0268	-0.0624	0.6608	0.6631
Change in Owner's Equity	0.1690	-112.1676	-0.0046	-0.1878	5.3877	799.5963

Table 4 – Descriptive statistics for one-year model variables[21]

Several ratios capture similar characteristics (see *Subsection 3.1.2*), e.g. *Wages/Revenues* and *Wages/Total Costs*, and high correlations can be expected between such ratios. *Table 5* indicates that correlations above 0.5 or below -0.5 are observed for 19 pairs of variables.

[21] The extreme mean and standard deviation of distressed clubs for *Change in Owner's Equity* is caused by Real Murcia CF SAD's value of -5757.1 due to a very low value of equity in the first year and a big loss that went into retained earnings in the second year. Hence, the median gives a much better indication of this variable.

	Spain	Population	Division (Y)	Division (Y-1)	Division Y-2	Relegation (Y-1)	Promotion (Y-1)	Short-term Movement (2 years)	Short-term Movement (3 years)	Long-term Movement (5 years)	Star Player (Y to Y-2)	Star Player (Y)	Star Player (Y-1)	Star Player (Y-2)	BV Equity / BV Liabilities	Wages / Revenues	Owner's Equity / Assets	Total Liabilities / Total Assets
Spain	1																	
Population	-0.0321	1																
Division (Y)	0.0057	-0.1620	1															
Division (Y-1)	-0.0082	-0.1230	**0.7637**	1														
Division Y-2	-0.0177	-0.1008	**0.7076**	**0.829**	1													
Relegation (Y-1)	0.0962	-0.1075	0.3402	-0.1682	-0.0521	1												
Promotion (Y-1)	0.0592	-0.0285	-0.113	0.4175	0.2721	-0.1454	1											
Short-term Movement (2 years)	0.1452	-0.1063	0.2918	0.3269	0.3205	0.4820	**0.5443**	1										
Short-term Movement (3 years)	0.1013	-0.0706	0.3592	0.3514	0.3473	0.4662	0.4430	**0.9117**	1									
Long-term Movement (5 years)	0.0943	-0.1341	0.4434	0.4516	0.4505	0.3379	0.3661	**0.744**	**0.8205**	1								
Star Player (Y to Y-2)	0.1128	0.3025	-0.3111	-0.2924	-0.3137	-0.0947	-0.0729	-0.1672	-0.1873	-0.2647	1							
Star Player (Y)	0.1160	0.3618	-0.2557	-0.2555	-0.2423	-0.1077	-0.1201	-0.2142	-0.2013	-0.2408	**0.7724**	1						
Star Player (Y-1)	0.0498	0.3033	-0.2485	-0.2715	-0.2575	-0.0623	-0.1276	-0.2019	-0.1969	-0.263	**0.8209**	**0.6973**	1					
Star Player (Y-2)	0.0869	0.2993	-0.2795	-0.2356	-0.2649	-0.1178	-0.0371	-0.1337	-0.1476	-0.2442	**0.8445**	**0.6766**	**0.6891**	1				
BV Equity / BV Liabilities	-0.0040	-0.0637	-0.0746	-0.0976	-0.0983	-0.0165	-0.0710	-0.0760	-0.0770	-0.0784	0.0012	-0.0216	-0.0093	0.0240	1			
Wages / Revenues	0.2673	-0.0173	0.0762	0.2385	0.1238	-0.0826	0.2800	0.1360	0.1042	0.1541	0.0389	-0.0048	-0.0863	-0.0149	-0.1344	1		
Owner's Equity / Assets	0.0978	-0.0925	-0.0483	-0.1577	-0.1206	0.0946	-0.1526	-0.0162	-0.0232	-0.0463	0.1126	0.0696	0.0903	0.1084	0.3917	-0.2026	1	
Total Liabilities / Total Assets	-0.0978	0.0925	0.0484	0.1578	0.1207	-0.0945	0.1526	0.0163	0.0233	0.0464	-0.1125	-0.0696	-0.0903	-0.1083	-0.3917	0.2027	-1	1
Wages / Total Costs	0.2575	0.0526	0.0748	0.1351	0.0462	-0.0047	0.1318	0.0909	0.0715	0.1020	0.0578	0.0184	-0.0063	0.0674	-0.0825	**0.5781**	-0.1316	0.1317
Amortization / Total Costs	0.2123	0.1193	-0.2575	-0.2632	-0.2648	0.0526	0.0150	0.0441	0.0561	-0.0871	0.3103	0.2216	0.1878	0.3232	-0.0521	0.1400	0.1290	-0.1289
Long-term Debt / Revenues	0.1915	0.1652	0.0475	0.0937	0.0733	-0.0410	0.0658	-0.0095	-0.0523	-0.0717	0.0387	0.0787	0.0636	0.0119	-0.1473	0.2235	-0.3081	0.3080
Revenues / Total Liabilities	-0.0886	-0.0848	0.1042	0.0253	0.1245	0.0891	-0.0793	0.0156	-0.0149	0.0639	-0.0673	-0.1070	-0.0318	-0.0710	0.3956	-0.2794	0.3049	-0.3048
Total Costs / Revenues	0.1009	-0.0770	0.0116	0.1705	0.0981	-0.1116	0.2405	0.1099	0.0937	0.0976	0.0145	-0.0296	-0.1061	-0.0404	-0.0542	**0.7137**	-0.1295	0.1297
Working capital / Total assets	0.1712	-0.0523	0.0681	0.0667	0.0404	0.0180	0.0194	0.0544	-0.0057	-0.0838	0.0001	-0.0306	-0.0363	0.0142	0.0289	0.0075	0.2153	-0.2153
Current Liabilities / Current Assets	-0.0525	0.0966	-0.0299	-0.0188	-0.0361	-0.0838	-0.0560	-0.1039	-0.0890	-0.0492	0.0008	0.0433	0.0020	-0.0106	-0.1618	0.1316	-0.2935	0.2935
Current Assets / Current Liabilities	-0.0368	-0.0280	0.0402	0.0380	0.0350	-0.0252	-0.0246	-0.0534	-0.0582	-0.0803	-0.0122	-0.0219	-0.0029	-0.0118	0.0324	-0.0472	0.0638	-0.0638
Cash assets / Current Liabilities	-0.0510	0.0231	-0.0824	-0.0669	-0.0567	0.0158	0.0412	-0.0459	-0.0862	-0.1050	0.1031	0.0530	0.1326	0.1014	0.2133	-0.1834	0.2360	-0.2360
Net Income / Total Assets	0.0536	0.0995	-0.1145	-0.2215	-0.1676	0.1030	-0.1450	-0.0352	-0.0642	-0.1358	0.1471	0.0744	0.1622	0.1186	0.1806	-0.3282	0.3876	-0.3877
EBIT / Total Assets	0.0810	0.0939	-0.1309	-0.2319	-0.1869	0.0888	-0.1467	-0.0374	-0.0503	-0.1433	0.1538	0.0914	0.1626	0.1376	0.1712	-0.4462	0.3195	-0.3196
Revenues / Total Assets	-0.1532	-0.0578	0.1789	0.1398	0.2193	0.0108	-0.0595	0.0118	0.0306	0.1387	-0.1092	-0.1344	-0.0640	-0.1285	-0.1394	-0.1874	-0.1940	0.1941
Change in Net Income	-0.0636	0.1346	-0.0070	-0.0302	0.1154	0.0418	-0.0119	-0.0202	-0.0004	-0.0452	0.1327	0.0815	0.1532	0.1431	-0.1135	-0.1103	-0.1304	0.1305
INTWO	-0.0644	0.0840	-0.0073	0.1093	0.1646	-0.0317	0.2214	0.0814	0.1062	0.0942	-0.0665	-0.0751	-0.1335	-0.0121	-0.1387	0.2641	-0.2081	0.2082
OENEG	-0.0107	0.0354	-0.0075	0.0495	0.0032	-0.1112	0.0198	-0.0079	-0.0154	-0.0417	-0.0480	-0.0384	-0.0613	-0.0721	-0.3567	0.3022	-0.5896	**0.5897**
Equity / Fixed Assets	0.0378	0.0107	-0.0169	-0.1103	-0.0868	0.0710	-0.1372	-0.0589	-0.0694	-0.0972	0.0498	0.0341	0.0397	0.0456	0.2281	-0.0947	**0.7735**	-0.7735
Change in Revenues	0.0205	-0.0064	-0.0079	-0.0608	0.3112	0.0711	-0.0490	0.1913	0.1673	0.1716	-0.0258	-0.0313	-0.0115	-0.0806	-0.0485	-0.1601	0.0044	-0.0043
Change in Owner's Equity	-0.1630	0.0130	0.0586	-0.0338	-0.0291	0.0244	-0.1717	-0.0577	-0.0301	-0.1182	0.0266	0.0204	0.0217	0.0225	0.0341	-0.3621	0.0242	-0.0243

Table 5 – Correlations between all variables tested; correlations higher than 0.5 and lower than -0.5 in bold

	Wages / Total Costs	Amortizat ion / Total Costs	Long-term Debt / Revenues	Revenues / Total Liabilities	Total Costs / Revenues	Working capital / Total assets	Current Liabilities / Current Assets	Current Assets / Current Liabilities	Cash assets / Current Liabilities	Net Income / Total Assets	EBIT / Total Assets	Revenues / Total Assets	Change in Net Income	INTWO	OENEG	Equity / Fixed Assets	Change in Revenues	Change in Owner's Equity
Wages / Total Costs	1																	
Amortization / Total Costs	0.2164	1																
Long-term Debt / Revenues	0.1321	0.1388	1															
Revenues / Total Liabilities	-0.0968	-0.3945	-0.3133	1														
Total Costs / Revenues	-0.1028	0.1093	0.1655	-0.2778	1													
Working capital / Total assets	0.1209	0.0658	0.0197	-0.0307	-0.094	1												
Current Liabilities / Current Asse	0.0294	-0.0122	0.0212	-0.2147	0.1311	-0.158	1											
Current Assets / Current Liabiliti	-0.0096	-0.0625	-0.0248	0.2192	-0.0534	0.0128	-0.0717	1										
Cash assets / Current Liabilities	-0.0390	-0.1651	-0.0617	0.4754	-0.2215	0.0236	-0.194	0.1456	1									
Net Income / Total Assets	0.1095	0.0933	-0.0787	0.1661	-0.5190	0.4067	-0.1597	-0.0085	0.2306	1								
EBIT / Total Assets	0.0516	0.1573	-0.0405	0.0753	-0.6029	0.3901	-0.1238	-0.0014	0.1953	**0.8562**	1							
Revenues / Total Assets	-0.0352	-0.4561	-0.2548	**0.6948**	-0.2301	-0.2441	-0.0616	0.1976	0.1961	-0.1652	-0.1992	1						
Change in Net Income	0.0634	0.0390	-0.0023	0.0259	-0.1879	0.0332	-0.0596	-0.0793	0.0486	0.2765	0.2008	0.1071	1					
INTWO	-0.0616	0.0701	0.1334	-0.2021	0.3942	-0.2242	0.0291	-0.0877	-0.1807	-0.4268	-0.3958	-0.0246	0.083	1				
OENEG	0.1087	-0.0426	0.2202	-0.3408	0.2425	-0.0886	0.3279	-0.0896	-0.2766	-0.3572	-0.3103	0.0518	0.0942	0.3064	1			
Equity / Fixed Assets	-0.1273	0.0632	-0.1180	0.2806	-0.0291	0.0758	-0.1608	0.0476	0.2382	0.1633	0.0937	-0.089	-0.1597	-0.0764	-0.364	1		
Change in Revenues	-0.1172	0.0335	-0.0554	0.2035	-0.0649	0.0079	-0.0689	-0.0058	0.0602	0.1510	0.1650	0.2001	0.3217	0.0090	-0.0729	-0.0468	1	
Change in Owner's Equity	-0.0788	-0.0763	-0.0468	0.0600	-0.2672	-0.0822	0.0117	0.0082	0.0393	0.0899	0.0735	0.0596	0.0195	-0.0816	-0.0942	0.0189	0.03	1

Table 5 (contd.) – Correlations between all variables tested; correlations higher than 0.5 and lower than -0.5 in bold

5.2 Model Determination

5.2.1 Principal Component Analysis

To decrease the number of correlated variables, principal component analysis is applied.

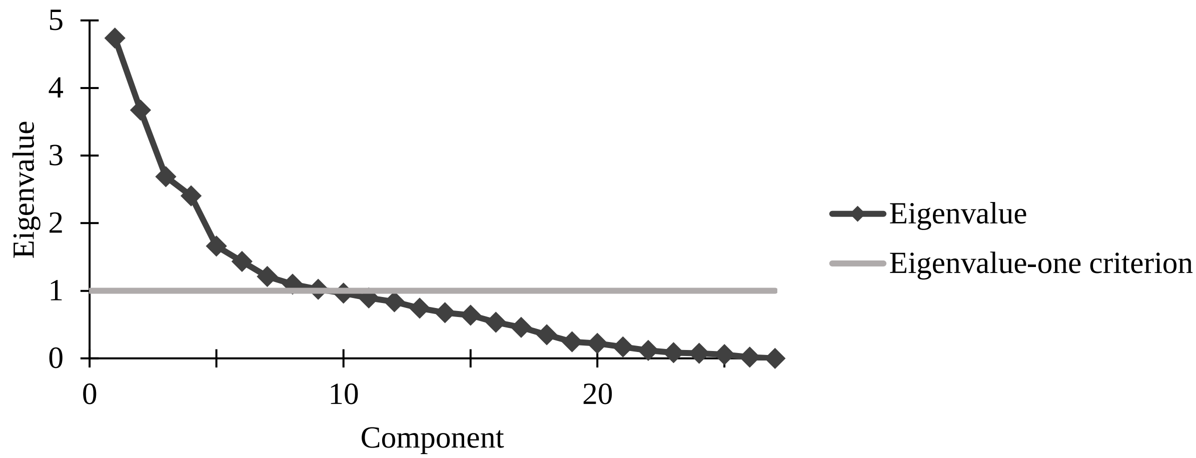


Figure 4 – Eigenvalues of components from PCA with eigenvalue-one criterion

Using the eigenvalue-one criterion demonstrates that nine components have eigenvalues above one (*Figure 4*). Hence, nine components are retained for analysis and the other components are dropped. To gain an understanding of what information the different components capture, analysis of the variable coefficients of each component is conducted. The promax rotation, an oblique rotation that allows for some correlation between the components, provides the simplest structure of the rotations. *Table 6* below shows the outcome of the promax rotation. Orthogonal rotations were also tried and they displayed similar patterns, but with less distinct coefficients for some of the variables.

Variable	Component 1	2	3	4	5	6	7	8	9
Population (01.01.2012)								0.7254	
Division Y			0.5513						
Division Y-1			0.5556						
Division Y-2			0.5438						
Short-term Movement (2 years)				0.5819					
Short-term Movement (3 years)				0.5949					
Long-term Movement (5 years)				0.5016					
BV Equity / BV Liabilities									
Wages / Revenues		-0.3413				0.5216			
Owner's Equity / Assets	0.5439								
Total Liabilities / Total Assets	-0.5439								
Wages / Total Costs						0.7796			
Amortization / Total Costs					-0.3056				
Long-term Debt / Revenues								0.3052	0.6625
Revenues / Total Liabilities					0.5479				
Total Costs / Revenues		-0.5209							
Working capital / Total assets								-0.4275	
Current Liabilities / Current Assets									-0.4322
Current Assets / Current Liabilities									0.4744
Cash assets / Current Liabilities					0.331				0.3507
Net Income / Total Assets		0.4534							
EBIT / Total Assets		0.4937							
Revenues / Total Assets					0.5624				
Change in Net Income							0.6023		
Equity / Fixed Assets	0.5597								
Change in Revenues							0.6877		
Change in Owner's Equity								0.3296	

Table 6 – Variable-component correlations above 0.3 (in absolute values) for the components with eigenvalues above 1

The principal component analysis underlines various patterns in the data. The highest variable-component correlation, in absolute values, ranges between 0.52 and 0.78. This is lower compared to the study of Skogsvik (1987), where the variable-component correlations ranged between 0.48 and 0.97, with the majority of the correlations attaining at least 0.70. However, the principal component analysis in the present study considers variables that are fundamentally different from each other whereas the PCA in Skogsvik's study analyzed variable groups that were highly homogeneous within each group (cf. Skogsvik, 1987, pp. 178-179 & pp. 183-185). Ideally, each variable should load highly on one component and low on the others as that would facilitate the distinction between components. However, the coefficients of the principal component analysis at hand provide sufficient information for recognizing patterns in the variables. *Table 7* provides the authors' interpretations of the nine retained components.

Component	**Content**
1	Capital structure
2	Profitability and cost structure
3	Division
4	Movements (relegations and promotions) in the last years
5	Asset turnover
6	Spending on players
7	Income development during last year
8	Access to funding (Buraimo, 2005)
9	Debt coverage

Table 7 – Interpretation of the components with eigenvalues above 1

5.2.2 Model Identification

The dummy variables identified in *Subsection 3.1.2* as well as one variable for each of the components identified in the previous subsection constitute a starting point to design the probit model. In order to drop insignificant variables and arrive at a more parsimonious model, a backwards elimination process is employed. In general, the variable that has the highest correlation to a certain component in the PCA is chosen to represent that specific component in the model. However, when a variable becomes insignificant at 20% significance level and there is a second variable with a similar correlation to the same component, the variable with the second highest correlation to the component is tested. The model that yields the highest pseudo R-squared and fulfills the significance criterion for each variable is presented in *Table 8*.

Pseudo R2 = 0.2047

Variable	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
Spain	0.7454	0.3073	2.43	0.02	0.143024	1.34773
Population	-2.20E-07	1.45E-07	-1.52	0.13	-5.04E-07	6.36E-08
Relegation (in Y-1)	0.8568	0.4035	2.12	0.03	0.0659777	1.647651
Shortterm Movement 2 years	-0.6301	0.2256	-2.79	0.01	-1.072395	-0.187879
StarPlayer (in Y to Y2)	-1.0295	0.4868	-2.11	0.03	-1.983535	-0.075407
Wages/Revenues	0.5627	0.3748	1.5	0.13	-0.171839	1.297163
Revenues/Total Liabilities	-0.3656	0.1224	-2.99	0.00	-0.605474	-0.125715
Equity/Fixed Assets	-0.0344	0.0235	-1.47	0.14	-0.080515	0.011625
_cons	-0.4541	0.3561	-1.28	0.20	-1.152146	0.243873

Table 8 – Statistical from the two-year model

As indicated by *Table 8*, all the included variables are significant at a 20% level. A 20% level was chosen to allow for a reasonable trade-off between the number of variables included in the model and the goodness-of-fit estimate, indicated by the Pseudo R-squared for logit/probit models (Wooldridge, 2012). The absolute values of the correlations between all variables are below 0.5 (*see Table 5*). The probit regression results in the following model for the determination of the probability of distress for a European football club which includes the constant and eight variables:

$$\begin{aligned} G = &-0.454 + 0.745 * Spain - (2.2 * 10^{-7}) * Population + 0.857 * Relegation\ (in\ Y-1) \\ &- 0.630 * Shortterm\ Movement\ 2\ years - 1.029 * Star\ Player\ (in\ Y\ to\ Y-2) \quad (6) \\ &+0.563 * Wages/Revenues - 0.366 * Revenues/Total\ Liabilities - 0.034 \\ &* Equity/Fixed\ Assets \end{aligned}$$

where G = An index value that can be looked up in the standard normal cumulative distribution function to give a probability of failure.

The coefficients of the variables *Spain*, *Population*, *Relegation*, *Star Player Y to Y-2*, *Wages/Revenues*, *Revenues/Total Liabilities*, and *Equity/Fixed Assets* are as expected (cf. *Appendix 1*). However, *Short-term Movement*, which indicates the number of relegations and promotions in the two years prior to distress, has a counter-intuitive coefficient sign. In a model of this kind, the coefficient and significance of a variable are also affected by the other variables. The descriptive statistics provide an indication as to the reason for this result. The survival clubs in the sample moved on average 0.48 times in the last two years whereas distressed clubs moved 0.40 times during the two years before distress. Similarly, a more detailed analysis indicates that survival clubs moved in 20.5% out of the total number of possible movements whereas distressed clubs only moved in 12.9% of the possible cases (*Table 9*). In line with expectations, survival clubs were promoted more often than relegated whereas distressed clubs had an equal amount of promotions and relegations. The coefficient of this variable is expected to be influenced by the characteristics of the sample.

	Number of Movements	% of possible movements	Promotions	Relegations
Survival	64	20.5%	40	24
Distress	16	12.9%	8	8

Table 9 – Comparison of the number of movements

Due to the fact that Spain has seen a large number of distressed clubs, the dummy variable *Spain* is highly significant. The legal prescription of not allowing penalization of Spanish football clubs for entering insolvency proceedings causes many clubs to seek such proceedings and consequently to be classified as distressed in the present study.

The significance of the variable *Population* gives indication that the size of the club's municipality is useful when predicting financial distress. It is based on the population in the municipality of the club on 01.01.2012 (Eurostat, 2015). These populations are assumed to remain reasonably constant during the period studied. In line with expectations, the coefficient of the variable shows that a higher population entails a lower risk of financial distress. In a larger municipality, the number of conceivable supporters is higher compared to in a smaller municipality. Thus, the potential revenue sources and access to funding is expected to be superior in larger municipalities (Buraimo, 2005).

In line with previous literature (Beech et al., 2010; Szymanski, 2012), *Relegation (in Y-1)* exhibits high statistical significance. It is a dummy variable that captures whether a club was relegated or not in the year before distress. As suggested by Beech et al. (2010), clubs tend to take excessive risks by investing heavily in players to ensure that they avoid relegation. Should they not prove successful and become relegated, the high amount of fixed costs cannot be covered by the lower revenue streams from less broadcasting revenues, lower ticket sales, and less merchandise sold (Ibid.). However, there is a second scenario that can help explaining why relegation is relevant. If a club is relegated, it might invest heavily to achieve direct promotion the subsequent season. Should this fail, the club is once again left with high fixed costs that are impossible to cover, which leads to financial distress.

The dummy variable *Star Player (in Y to Y-2)* is also highly significant. It captures whether a club has had a star player during the current or past two seasons. The coefficient of the variable indicates that a club with a star player is less likely to fail, which is in line with expectations. Despite resulting in higher costs and requiring high transfer fees, star players should generate an increase in revenues through improved on-field performance as well as through higher ticket and merchandise sales. In addition to the aforementioned, a star player provides a buffer in case of distress. When approaching financial distress, a club can sell this valuable asset, increase solvency and cover debts with the proceeds.

Wages/Revenues is, as suggested by previous research (Barajas & Rodríguez, 2010; Szymanski, 2012), relevant to the model. On average, wages correspond to roughly 63% of the revenues for European clubs, and consequently constitute the bulk of total costs (UEFA, 2017). Previous literature has found that when wages are close to, or exceeding, the revenues of a club, an increase of indebtedness is likely to follow, which, in turn, increases the risk of financial distress (Barajas & Rodríguez, 2010; Mourao, 2012; Szymanski, 2012). Hence, the coefficient and significance of *Wages/Revenues* are in line with expectations.

Revenues/Total Liabilities can be thought of as the extent to which a club is able to cover its liabilities with the revenues generated in one year. Intuitively, a higher ratio entails a lower probability of distress, which is confirmed by the sign of the coefficient. Szymanski (2012) found this ratio to be related to *Wages/Revenues*. Specifically, he argued that a decreasing *Revenues/Total Liabilities* ratio is a long-term result indicating the "deterioration of the balance sheet" (Szymanski, 2012, p. 15), caused by excessive wages being covered by taking on additional liabilities.

Finally, *Total Equity/Fixed Assets* is the variable with the lowest significance of all the variables included in the model. Nevertheless, it provides information on how fixed assets of a club are financed. Marcinkowska (2013) found that equity should at least cover the fixed assets for a club to be financially healthy. However, as demonstrated previously, there are many clubs with negative equity that are not in a distressed situation, which could help explaining the relatively lower significance of the variable.

The model development process is successful in determining a model that can be expected to be able to differentiate between distressed and survival clubs based on the available data for European football clubs. However, the relevance and validity of the model are contingent on the accuracy with which it can predict financial distress. Hence, the next subsection tests the accuracy of the model on the sample of analysis as well as on a hold-out sample achieved through the jackknife procedure.

5.3 Accuracies

The accuracy in the present study is measured in terms of arithmetic average error rate ($rate(\bar{e})$) and weighted average error rate ($rate(\bar{e})'$), as given by *Equation 2* and *3* in *Subsection 3.2.2*. There are two ways of putting these error rates into context, the first of which is to compare against the accuracies achieved in previous studies by referring to *Table 1* in *Subsection 2.2*. The second is to compare the error rates against those produced from naïve decision rules. As an additional way of contextualizing the error rates achieved by the model developed in this study, an application of Zmijewski's (1984) model, found by Gerritsen (2015) to be the best applicable bankruptcy prediction model on football clubs, on the data sample used in this study follows in *Subsection 6.1*. To be consistent with previous research where one-to-one matched samples of failed and surviving firms have been prominent, $rate(\bar{e})$ is used for comparing the results to other studies. For comparing the results against the naïve decision rules, $rate(\bar{e})'$ is used.

	$\boldsymbol{rate(\bar{e})}$	$\boldsymbol{rate(\bar{e})'}$
Sample of Analysis	28.2%	19.2%
Hold-out Sample	32.1%	23.1%

Table 10 – Model accuracy for a one-year prediction horizon

As *Table 10* shows, both $rate(\bar{e})$ and $rate(\bar{e})'$ are slightly higher for the hold-out sample compared to the sample of analysis, which is expected given that there is no overlap between the clubs used for estimating the model and those classified in the hold-out case. Also in line with expectations, $rate(\bar{e})$ is higher than $rate(\bar{e})'$ for both the sample of analysis and the hold-out sample.

Vis-à-vis the results for the research presented in *Subsection 2.2*, a $rate(\bar{e})$ of 32.1% for a one-year prediction on a hold-out sample is relatively high. Several studies showed $rate(\bar{e})$ in the vicinity of

20% or lower (e.g. Beaver, 1966; Altman, 1968; Ohlson, 1980; Skogsvik, 1987; Altman, 2000; Gu, 2002). Worth acknowledging is that Ohlson (1980) did not test predictive accuracy on a hold-out sample but instead only examined the sample of analysis. While the error rates in the present study are high compared to some previous studies, the model of Zavgren (1985) exhibited a similar $rate(\bar{e})$ of 31% for one-year predictions. Furthermore, the accuracy of the model in this study is slightly better compared to Gerritsen's (2015) application of Zmijewski's (1984) model, where 34% of the clubs were misclassified. However, this is not $rate(\bar{e})$ but rather $rate(\bar{e})'$ with the proportions of that specific sample and as such the accuracy provided by Gerritsen is not directly comparable to the other studies.

The relatively low prediction accuracy achieved by the model developed in the present study compared to previous prediction models can be explained by the distinct characteristics of the football industry. As presented in *Subsection 2.1* and shown in *Subsection 5.1*, clubs tend to take on excessive risks and make losses. These are not features unique to distressed clubs, given that the descriptive statistics indicate that even the surviving clubs in the sample make losses on average. As a result, the distinction between distressed and surviving becomes difficult for the model to make, leading to a relatively low accuracy rate.

However, the error rates from the model developed in this study are significantly lower compared to those that would have ensued from applying a naïve decision rule. The expected $rate(\bar{e})'$ from an application of Rule 2, established in *Subsection 3.2.2* as a decision rule that classifies clubs randomly but proportionally based on the a priori probabilities of the sample, would equal 37.5%[22]. Comparing the $rate(\bar{e})'$ of 23.1% for the hold-out sample to the naïve decision rule shows that the $rate(\bar{e})'$ of the model is statistically lower at a significance level of 1%[23]. Hence, despite a low accuracy rate compared to previous studies, the model does add value by significantly outperforming the naïve decision rule.

Figure 5 graphically illustrates the trade-off between type I and type II errors in the one-year model. At the lowest cut-off value possible, only type II errors are made and at the highest cut-off value possible, only type I errors are made. By varying the critical cut-off value, fewer type II errors are made at the expense of making more type I errors and vice versa.

---

[22] $Expected\ prediction\ error = Proportion\ of\ distressed\ clubs * Probability\ of\ type\ I\ error + Proportion\ of\ survival\ clubs * Probability\ of\ type\ II\ error = \frac{1}{4} * \frac{3}{4} + \frac{3}{4} * \frac{1}{4} = 37.5\%$

[23] $z = \frac{(\hat{p} - p_0)}{\sqrt{\frac{p_0 * (1-p_0)}{n}}} = \frac{23.1\% - 37.5\%}{\sqrt{\frac{37.5\% * (1-37.5\%)}{208}}} = -4.3$ p-value < 0.01 (Sharpe, De Veaux, Velleman, 2012)

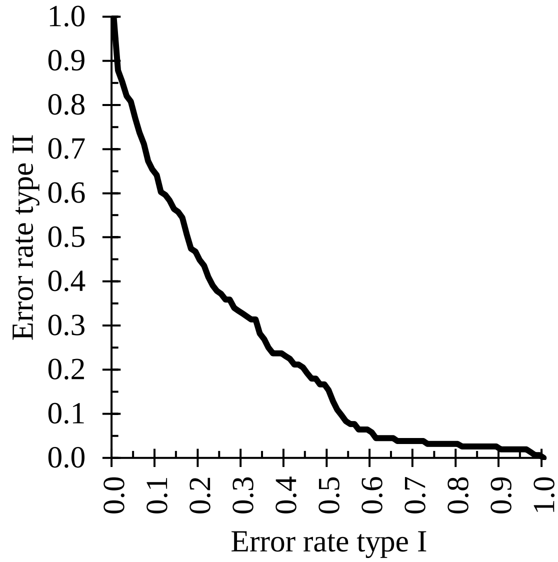


Figure 5 – Trade-off between type I & type II error rates for the hold-out sample

5.4 Two-year Model

5.4.1 Model Identification

Given that the one-year prediction model outperformed the naïve decision rule, estimating the two-year prediction model is worthwhile. Previous bankruptcy prediction studies found that model accuracies generally tend to deteriorate the further away in time from the distress event the prediction takes place (cf. Bellovary et al., 2007, pp. 23-41). Hence, the authors of this study expect the accuracies of the two-year model to be lower compared to the accuracies presented for the one-year model.

The methodology used to derive the model and examine its accuracies is the same as for the one-year model. Since the two-year model utilizes data two years prior to distress, it requires that data for the clubs is available for an additional year compared to the one-year model. This leads to a decrease in sample size where 49 distressed clubs are identified, compared to 56 for the one-year sample. These 49 distressed clubs are matched against two survival clubs each, entailing that the two-year sample consists of 147 clubs in total. Given that the sample would have decreased further if variables that require financial information for two consecutive years (i.e. OENEG, INTWO) had been included, these variables are no longer considered in the model. None of these variables was significant in the one-year model and, as such, the development of the two-year model is not expected to be significantly impacted by this decision. At the same time, it is important to acknowledge the risk of omitted-variable bias from not considering these variables. In total, 31 variables are examined for developing the two-year model. The descriptive statistics for the two-year model are found in *Appendix 3*.

Tests for correlations show that 29 correlations (in absolute values) are above 0.5 (*Appendix 4)*. Consequently, principal component analysis is employed in order to reduce the number of variables. The results of the PCA show that seven components have eigenvalues above one (*Appendix 5*). As for the one-year model, the promax rotation provides the simplest structure where the variables have the highest loadings on each component. The variable coefficients of the components and the interpretations of the components are included in the appendix (*Appendix 6*; *Appendix 7*). Again, ideally the coefficients of the variables on the components could have been higher to facilitate distinctions, but patterns can nonetheless be identified. Several components are similar to the ones identified for the one-year model.

Following the principal component analysis, the backwards elimination process was applied and resulted in the following statistical output from the two-year model:

Pseudo R2 = 0.1548

Variable	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
Spain	1.0906	0.3201	3.41	0.00	0.4631603	1.718047
Shortterm Movement 2 years	-0.3078	0.1915	-1.61	0.11	-0.68319	0.067503
StarPlayer (Y-1 to Y-3)	-1.0213	0.4460	-2.29	0.02	-1.895457	-0.147225
Total Liabilities/Total Assets	0.1685	0.1253	1.34	0.18	-0.077116	0.414046
Cash Assets/Current Liabilities	-1.2435	0.5887	-2.11	0.04	-2.39735	-0.089704
_cons	-0.4144	0.2245	-1.85	0.07	-0.85446	0.025598

Table 11 – Statistical output from the two-year model

The absolute values of the correlations between these variables are again below 0.5 (*Appendix 4*). The pseudo R-squared is lower than for the one-year model, which is in line with expectations. At a 20% significance level only five variables were found to be significant. Provided the statistical output in *Table 11*, *Equation 7* gives the two-year prediction model for financial distress of European football clubs:

$$
\begin{aligned}
G = -0.414 + 1.091 * Spain - 0.308 * Shortterm\ Movement\ 2\ years \\
- 1.021 * StarPlayer\ (Y - 1\ to\ Y - 3) + 0.168 * Total\ Liabilities/Total\ Assets \\
- 1.244 * Cash\ Assets/Current\ Liabilities
\end{aligned}
\tag{7}
$$

where G = An index value that can be looked up in the standard normal cumulative distribution function to give a probability of failure.

All variables except *Short-term Movement 2 Years* are in line with expectations (cf. *Appendix 1*). The reason for *Short-term Movement 2 Years* being negative is the same as for the one-year model. However, the significance of the variable is lower in the two-year model. *Spain, Short-term Movement 2 Years* and *Star Player Y-1 to Y-3* have the same interpretations as for the one-year model. *Total Liabilities/Total Assets* and *Cash Assets/Current Liabilities* are variables that were not part of the one-year model.

The ratio *Total Liabilities/Total Assets* represents the capital structure of the club, and more specifically the amount of assets relative to the amount of liabilities. It is a financial variable derived from both previous bankruptcy prediction (Beaver, 1966; Ohlson, 1980) as well as football literature (Barajas & Rodríguez, 2010; Mourao, 2012; Szymanski, 2012). If the value exceeds one, the club has negative equity. Szymanski (2012) found this value to increase when clubs have excessive player wages while simultaneously the on-field performance is lower than expected, negatively impacting revenues. In line with expectations, the coefficient of *Total Liabilities/Total Assets* is positive, indicating that higher liabilities in relation to assets increase the risk of distress.

Cash Assets/Current Liabilities shows the extent to which a club is able to cover all of the current liabilities with the amount of cash it has readily available (Skogsvik, 1987). It is a measure of financial health that is expected to be higher for financially non-distressed clubs, which is indicated by the sign of the coefficient. However, given that it could be argued that liquid assets such as cash are relevant for short-term liquidity, it is interesting observing this variable's significance in the two-year model and not in the one-year model.

<u>5.4.2 Two-year Prediction Accuracies</u>

	$\boldsymbol{rate(\bar{e})}$	$\boldsymbol{rate(\bar{e})'}$
Sample of Analysis	33.2%	20.4%
Hold-out Sample	31.6%	27.2%

Table 12 – Prediction accuracies for the two-year model

To some extent, the prediction accuracies for the two-year model are in line with expectations. Interestingly, however, $rate(\bar{e})$ for the hold-out sample is slightly lower for a prediction horizon of two years compared to a prediction horizon of one year. Furthermore, for the two-year model $rate(\bar{e})$ is in fact lower for the hold-out sample than for the sample of analysis. While these findings are unexpected, $rate(\bar{e})$ is higher for the classification of the sample of analysis and $rate(\bar{e})'$ is higher for both the sample of analysis and the hold-out sample in the two-year model compared to the one-year model.

As mentioned earlier, error rates are expected to increase when extending the forecasting period. This is not observed for $rate(\bar{e})$ when testing the model's accuracy on the hold-out sample in this study. A $rate(\bar{e})$ of 31.6% for a two-year prediction is still significantly higher compared to the studies listed in *Subsection 2.2* that made two-year predictions on hold-out samples but the magnitude of the difference to those studies has been reduced. The lower performance of the present model can again be explained by the unique characteristics of the football industry presented in *Subsection 2.*1. Nevertheless, the two-year prediction $rate(\bar{e})$ is close to that of Zavgren's (1985) model, where a $rate(\bar{e})$ of 31% was observed for a forecast horizon of two years. Gerritsen (2015) found that Zmijewski's (1984) model had a $rate(\bar{e})'$ of 37% when applied to Dutch clubs two years before distress, which is comparably higher than the 31.6% achieved by the model developed in this study. Again, given that Gerritsen does not provide $rate(\bar{e})$, the results are not directly comparable. Lastly, Altman's (1968) model was only tested on a hold-out sample for a one-year forecast, but when applied to the sample of analysis for a two-year prediction, the model achieved a $rate(\bar{e})$ of 28%. This is comparable to the $rate(\bar{e})$ of the model in this study, given that the errors are expected to rise if his model was applied to a hold-out sample for the two-year forecast.

$rate(\bar{e})'$ has increased from 23.1% to 27.2% in the two-year prediction for the hold-out sample. However, the a priori probability of failure has simultaneously increased from 25% to 33% due to the changes in the sample proportion of financially distressed and surviving clubs. Consequently, the $rate(\bar{e})'$ for the naïve decision rule is now roughly 44% when randomly classifying one third of the clubs as financially distressed and two thirds of the clubs as surviving. The difference in error proportions is again significant at a 1% level[24] and it can be concluded that the two-year model also outperforms the naïve decision rule when applied to the hold-out sample of this study.

The fact that all error rates except the minimum $rate(\bar{e})'$ on the hold-out sample are relatively similar compared to the one-year model indicates that timing is not as relevant of an issue in the prediction of financial distress for football clubs as in other industries. Given that the error rates are similar for one and two-year prediction horizons, it seems that the model is able to differentiate between distress and survival but that predicting the timing of distress has less of an effect. One possible reason is that clubs can sustain a long period in a bad financial condition but eventually do fall into the distressed category as defined in this study.

Figure 6 graphically illustrates the trade-off between type I and type II errors on the hold-out sample in the two-year model. Compared to *Figure 5* that shows the corresponding trade-offs for the one-year model, the shape of *Figure 6* is more linear, which indicates that the two-year model is worse at distinguishing between distressed and surviving clubs. Thus, even though the minimum error rates are similar, *Figure 5* and *Figure 6* exhibit that the one-year model generally has a better trade-off between type I and type II errors.

Even though the accuracy of the two-year model significantly outperforms the naïve decision rule, a three-year model is not developed in this study. The main reason is that additional clubs would have been dropped due to data availability reasons and the sample would have been unjustifiably small. Furthermore, *Figure 6* indicates that the two-year model is not as accurate in distinguishing between distressed and surviving clubs as the one-year model. A three-year model estimated on fewer clubs is expected to be even less accurate.

[24] $z = \frac{27{,}2\% - 44{,}4\%}{\sqrt{\frac{44{,}4\% * (1 - 44{,}4\%)}{156}}} = -4{,}3$ p-value < 0.01

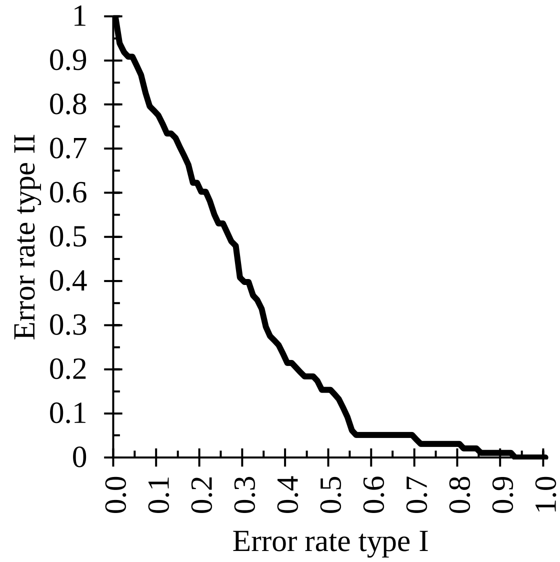


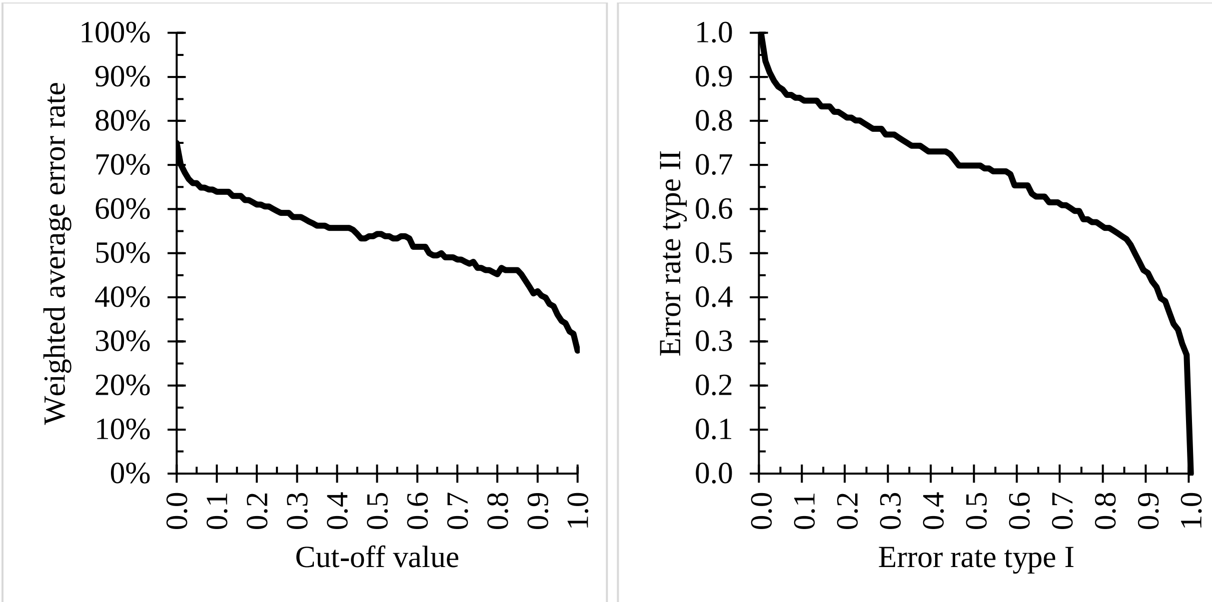
Figure 6 – Trade-off between the error rates for the two-year model on the hold-out sample

6. Additional Tests

This section presents the results of applying the Zmijewski model (1984) to the sample used in this study. Thereafter, a discussion on the calibration formula required to get unbiased probabilities of distress follows.

6.1 Application of Zmijewski's (1984) Model

As introduced in *Subsection 2.3*, Gerritsen (2015) found Zmijewski's (1984) model to be the best applicable bankruptcy prediction model for Dutch football clubs in the 2010-2014 period. The model misclassified 34-39% of the clubs depending on the prediction period at a cut-off value of 0. Hence, Gerritsen suggested that Zmijewski's (1984) model was more relevant to the football industry compared to Altman's (2000) and Ohlson's (1980) models. Applying Zmijewski's (1984) model to the data sample used in this study enables a comparison of the accuracies achieved by the model developed in this study and a more traditional bankruptcy prediction model when applied to football clubs.



Figures 7 & 8 – Weighted average error rate for the Zmijewski model with varying cut-off values and the trade-off between type I and type II error rates

As indicated in *Figures 7* and *8*, the cut-off value that minimizes $rate(\bar{e})'$ can be found above the 99th percentile. At that cut-off value, the type I error rate is maximized and the type II error rate is minimized. At lower percentiles, the type II error rate is relatively high, indicating that the Zmijewski model misclassifies surviving clubs as distressed. Using Gerritsen's cut-off value of zero, which is at the 50th percentile, 54.3% of the clubs are misclassified. This is significantly worse compared to the naïve decision rule of 37.5%, which is in line with expectations given that Zmijewski's model was estimated on a set of fundamentally different, profit-maximizing firms over 30 years ago. As discussed in *Subsection 2.1*, even non-distressed football clubs can exhibit negative equity and consecutive years of net-losses. Thus, when applying traditional bankruptcy prediction models to the European football industry, many of these clubs are expected to be misclassified as distressed instead of surviving. This is evidently the case when applying Zmijewski's model to the sample of this study.

Using a cut-off value close to the 100th percentile for the Zmijewski model, the minimum percentage of $rate(\bar{e})'$ is 27.9% (*Figure 7*). At this extreme cut-off value, however, 73.1% of the distressed clubs are misclassified and 12.8% of the surviving clubs are still classified as distressed. Compared to the $rate(\bar{e})'$ of the one-year model in this study of 23.1%, 27.9% is relatively higher but nevertheless lower than the naïve decision rule at 37.5%. Hence, the Zmijewski model also outperforms the naïve decision rule at an extreme cut-off value but its accuracy is lower than that achieved by the model developed in this study.

6.2 Calibration

As discussed in *Subsection 3.3*, the probabilities derived from the models presented in this study are biased by the sample proportions of distressed and surviving clubs and need to be calibrated in order to be used in real-world decision settings. In this particular study, the a priori probability of distress in the sample of clubs is 0.25 for the one-year model and 0.33 for the two-year model. To apply the calibration formula presented by Skogsvik and Skogsvik (2013) (given by *Equation 5*), the a priori probability of distress in the population of European football clubs is required.

Previous literature gives indications as to what the a priori probability of distress of football clubs in various European countries can be. Scelles et al. (2016) observed that the frequency of insolvency proceedings per season in France between 1970 and 2014 was 3% and Szymanski (2012) found the corresponding frequency in England between 1974 and 2010 to be 2%. The authors of the present study compiled all cases of financial distress that have occurred in the countries and divisions examined in this study during the period 2006-2016. This process led to an identification of 112 cases of financial distress out of 4042 club seasons in Europe, resulting in an a priori probability of financial distress of 2.8%. The estimation of 2.8% lies between the 3% suggested for France (Scelles et al., 2016) and the 2% observed in England (Szymanski, 2012). Hence, the authors of this study suggest using an a priori probability of 2-3% for the population when calibrating the sample-biased probabilities generated by the model developed in this study. Even though Szymanski (2010) found that the financial crisis had little impact on the probability of financial distress in England, it is not unreasonable to assume that external economic and regulatory factors can have an effect on the a priori probability of financial distress for the population.

7. Concluding Remarks

The aim of this study was to develop a financial distress prediction model for European football clubs that makes use of publicly available data. The development of such a model is successful and, in line with expectations, financial as well as non-financial variables that are relevant specifically to football clubs in combination with financial variables that are prominent in the bankruptcy prediction literature are used. The predictive accuracy for both the one-year model as well as the two-year model is significantly better than the accuracy given by applying naïve decision rules. However, compared to prominent bankruptcy prediction models, the accuracy achieved by the model developed in the present study is generally lower. The lower prediction accuracy is attributed to the unique conditions inherent in the football industry, e.g. that even clubs classified as surviving often exhibit features such as negative equity, which obfuscates the distinction between distressed and survival clubs. Furthermore, the definition of financial distress in this study was chosen at an early stage of a club's financial struggles. Consequently, the accuracy rates are expected to be lower compared to studies that define the initiation of insolvency proceedings or the actual filing for bankruptcy as the earliest definitions of distress. At the later stages of a company's financial struggles, the differences between the financial ratios of distressed and survival clubs are expected to be more pronounced. Lastly, the forecast horizon for the one-year prediction model is, on average, 18.8 months, which is significantly longer compared to other studies (e.g. Lennox, 1999; Skogsvik, 1987).

Thus, the model has predictive power but the achieved accuracy should not be overemphasized. However, application of Zmijewski's (1984) model, found by Gerritsen (2015) to be the bankruptcy prediction model with the highest predictive accuracy when applied to football clubs, shows that the model developed in the present study is superior in its ability to distinguish between distressed and non-distressed clubs. Thus, football club stakeholders assessing the survival likelihood of a club within a similar institutional context to the sample of this study are better off using the model developed in this study compared to other prediction models or naïve decision rules. However, there are limitations related to the practical applicability of the model that arise from the non-random nature of the sample in this study. These limitations, discussed in detail in the following subsection, are important to acknowledge when implementing the model.

7.1 Limitations

The distressed and survival clubs examined in this study are not randomly selected. The 208 clubs included constitute the entire population of clubs for which information is available. The non-random sample entails that inferences are limited to the clubs in the sample and that inferences on all European football clubs cannot be made. Consequently, the extent to which the model developed in this study can be applied outside the sample is questionable. However, given that the economic, legal and political environments where the model is applied are similar to the settings under which the model was developed, the model can be expected to work reasonably well. Therefore, application of the model on European clubs that are not included in the sample is likely to be better than using the model for, e.g., American clubs that cannot be relegated from a division and therefore are expected to have a very different business model.

A second limitation of this study is regarding the identification process of distressed clubs. Given that there is no systematic method of identifying distressed clubs in Europe, there is a risk that the process of identifying these clubs was subject to errors by the authors of this study, both in terms of interpretations of the information as well as the possibility that some distressed clubs were not identified. Furthermore, this study assumes that similarities exist between the different countries and football associations included in the sample. Even though Spain was found to be significantly different by not allowing penalization of clubs for entering insolvency proceedings, there may be other important differences between the countries in terms of legal systems, licensing systems as well as general business environments that were not identified by the authors of the present study. Clubs also have different legal statuses, which was not considered in detail by the authors of this study.

There are further ratios and indicators that would have been desirable to include in the model. However, due to data unavailability, as was discussed in *Subsection 4.2,* several indicators that potentially could have been relevant in predicting financial distress for football clubs were not considered. Ratios that build on information from the cash-flow statements are examples of indicators that were found relevant in previous research (Marcinkowska, 2013), but that are not considered in the present study due to the unavailability of such statements in the database used. For certain non-financial indicators, the necessary information was not publicly available or impossible to collect systematically. An example of a non-financial indicator that potentially could have been useful is average attendance during a club's home games. Including such a ratio could have shed light on the relationship between ticket revenue and financial distress, which was found to be significant in earlier studies (Scelles et al., 2016; Szymanski, 2012). Also related to data availability is the fact that the hold-out sample was created through the jackknife procedure. This entails that there is only cross-sectional validation of the models and no time-series validation. Consequently, the question how models developed in this study stand the test of time remains unanswered.

Barajas and Rodríguez (2010) suggested that the quality of the financial information provided by Spanish football clubs was of low quality. They recognized a number "of unqualified opinions in the audit reports" (Barajas and Rodríguez, 2010, p. 54) as well as a lack of accounting quality for several items of the financial statements. Inferences made on the basis of such low-quality information have to be interpreted cautiously. Lastly, although outliers were kept since they arguably contain information about the underlying sample, winsorizing the data could have mitigated their impact.

7.2 Suggestions for Future Research

In line with the limitations of this study, most of the suggestions for future research are related to data. The authors suggest that future research examines cash flow-related ratios, given that previous bankruptcy prediction studies (cf. Bellovary et al., 2007, p. 42) and football literature (Marcinkowska, 2013) have found these types of ratios to be relevant predictors of financial distress. However, this would require new means of accessing data. Such means, to the best of the authors' knowledge, do not yet exist. Linked to data accessibility is also the suggestion to study *all* European football clubs. Furthermore, including spectator numbers from clubs' home games as a variable is expected to be useful to predict financial distress (Scelles et al., 2016; Szymanski, 2012).

The present study was limited by the number of years available in the Amadeus database. As such, the sample size decreased significantly for each additional forecasting year. By new means of accessing data, future research can extend the forecasting period to examine if, as suggested by the results in the present study, prediction of financial distress for football clubs is less affected by temporal aspects. By successfully extending the forecasting horizon, users of the model can make informed decisions about clubs and act to prevent distress long before it occurs.

References

Literature

Abdi, H. & Williams, L.J. (2010). ‘Principal Component Analysis’. *Wiley Interdisciplinary Reviews: Computational Statistics*, *2*(4), 433-459.

Agarwal, V. & Taffler, R.J. (2007). ‘Comparing the Performance of Market-based and Accounting-based Bankruptcy Prediction Models’. *Journal of Banking and Finance*, *32*(8), 1541-1551.

Altman, E. (1968). ‘Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy’. *The Journal of Finance*. *23*(4), 589-609.

Altman, E. (2000). ‘Predicting Financial Distress of Companies: Revisiting the Z-Score and ZETA® Models’. *Stern School of Business, New York University*, 9-21.

Ascari, G. & Gagnepain, P. (2006). ‘Spanish Football’. *Journal of Sports Economics*, *7*(1), 76-89.

Barajas, Á. & Rodríguez, P. (2010). ‘Spanish Football Clubs' Finances: Crisis and Player Salaries’. *International Journal of Sport Finance, 5*(1), 52-66.

Barajas, Á. & Rodríguez, P. (2014). ‘Spanish Football in Need of Financial Therapy: Cut Expenses and Inject Capital’. *International Journal of Sport Finance, 9*(1), 73-90.

Beaver, W. H. (1966). ‘Financial Ratios as Predictors of Failure’. *Journal of Accounting Research, 4*, 71-111.

Beech, J., Horsman, S. & Magraw, J. (2010). ‘Insolvency Events among English Football Clubs’. *International Journal of Sports Marketing and Sponsorship*, *11*(3), 53-66.

Bellovary, J. L., Giacomino, D. E. & Akers, M. D. (2007). ‘A Review of Bankruptcy Prediction Studies: 1930 to Present’. *Journal of Financial Education*, 33, 1-42.

Boeri, T. & Severgnini, B. (2014). ‘The Decline of Professional Football in Italy’. *Handbook on the Economics of Professional Football*, 322-335.

Boscá, J. E., Liern, V., Martínez, A. & Sala, R. (2008). 'The Spanish Football Crisis'. *European Sport Management Quarterly*, *8*(2), 165-177.

Buraimo, B., Forrest, D. & Simmons, R. (2007). 'Freedom of Entry, Market Size, and Competitive Outcome: Evidence from English Soccer'. *Southern Economic Journal*, *74*(1), 204-213.

Cooper, C. & Joyce, Y. (2013). 'Insolvency Practice in the Field of Football'. *Accounting, Organizations and Society*, *38*(2), 108-129.

Dey, E.L. & Astin, A.W. 1993. 'Statistical Alternatives for Studying College Student Retention: A Comparative Analysis of Logit, Probit, and Linear Regression'. *Research in Higher Education*, *34*(5), 569-581.

Dietl, H. M. & Franck, E. (2007). 'Governance Failure and Financial Crisis in German Football'. *Journal of Sports Economics*, *8*(6), 662-669.

Faraway, J. (2014). *Linear Models with R, 2nd Edition.* New York, USA: Chapman and Hall.

Grice J & Ingram R. (2001). 'Tests of the Generalizability of Altman's Bankruptcy Prediction Model'. *Journal of Business Research, 54*(1), 53-61.

Gerritsen, P. L. (2015). *Accuracy Rate of Bankruptcy Prediction Models for the Dutch Professional Football Industry* (Master's Thesis, University of Twente).

Georgievski, B. (2016). 'Is English Football Facing a Financial Crisis, or are We Only Missing Better Regulation?'. *Journal of Tourism, Hospitality and Sports, 18,* 1-9.

Karels, G. V. & Prakash, A. J. (1987). 'Multivariate Normality and Forecasting of Business Bankruptcy'. *Journal of Business Finance & Accounting*, *14*(4), 573-593.

Hembert, E., Lucero, M. G. R., Mesnard, X., Rossi, L. & Rothenbücher, J. (2010). 'The A.T. Kearney EU Football Sustainability Study: Is European Football Too Popular to Fail?'. *A.T. Kearney, Inc*. 1-8.

Kuper, S. & Szymanski, S. (2014). *Soccernomics*. New York, USA: Harpersport.

Kim, H. & Gu, Z. 2006. 'Predicting Restaurant Bankruptcy: A Logit Model in Comparison with a Discriminant Model'. *Journal of Hospitality and Tourism Research*, *30*(4), 474-493.

Lennox, C. (1999). ‘Identifying Failing Companies: A Re-evaluation of the Logit, Probit and DA Approaches’. *Journal of Economics and Business, 51*(4), 347-364.

Marcinkowska, E. (2014). ‘KSP Polonia Warszawa's Financial Problems - Analysis of Liquidity and Debt’. *Journal of Entrepreneurship Management and Innovation, 9*(1), 135-155.

Margret, J. E. (2002). ‘Insolvency and Tests of Insolvency: An Analysis of the “Balance Sheet” and “Cashflow” Tests’. *Australian Accounting Review, 12*(27), 59-72.

Morrow, S. (2014). ‘Football Finances’. *Handbook on the Economics of Professional Football*, 80-99.

Mourao, P. (2012). The Indebtedness of Portuguese Soccer Teams - Looking for Determinants’. *Journal of Sports Sciences, 30*(10), 1025-1035.

O’Rourke, N., Hatcher, L. & Stepanski, E. J (2005). *A Step-by-step Approach to Using SAS for Univariate and Multivariate Statistics (2nd ed.).* New York, USA: John Wiley.

Niitsuma, H. & Okada, T. (2005). ‘Covariance and PCA for Categorical Variables’. In: *Proceedings of the 9th Pacific-Asia Conference on Advances in Knowledge Discovery and Data Mining*, 523-528. Berlin, Germany: Springer-Verlag.

Ohlson, J. A. (1980). Financial Ratios and the Probabilistic Prediction of Bankruptcy. *Journal of Accounting Research*, *18*(1), 109-131.

Scelles, N., Szymanski, S. & Dermit-Richard, N. (2016). ‘Insolvency in French Soccer: The Case of Payment Failure’. *Journal of Sports Economics,* epub ahead of print.

Sharpe, N., De Veaux, R. D. & Velleman, P. (2012). *Business Statistics*. Boston, USA: Pearson Education, Inc.

Skogsvik, K. (1987). *Prognos av Finansiell Kris med Redovisningsmått: En Jämförelse Mellan Traditionell och Inflationsjusterad Redovisning*. Stockholm, Sweden: Ekonomiska Forskningsinstitutet vid Handelshögsskolan i Stockholm.

Skogsvik, K. (1990). ‘Current Cost Accounting Ratios as Predictors of Business Failure: The Swedish Case’. *Journal of Business Finance & Accounting*, *17*(1), 137-160.

Skogsvik, K. & Skogsvik, S. (2013). 'On the Choice Based Sample Bias in Probabilistic Bankruptcy Prediction'. *Investment Management and Financial Innovations, 10*(1), 29-37.

Smith, T. (1983). 'On the Validity of Inferences from Non-random Sample'. *Journal of the Royal Statistical Society. Series A (General), 146*(4), 394-403.

Szymanski, S. (2010). 'The Financial Crisis and English Football: The Dog That Will Not Bark'. *International Journal of Sport Finance, 5*(1), 28-40.

Szymanski, S. (2012). 'Insolvency in English Professional Football: Irrational Exuberance or Negative Shocks?'. *International Association of Sports Economists, Working Paper Series 1202*.

The Law Firm Network (2013). 'International Insolvency and Restructuring Guide'. *The Law Firm Network and Contributors*.

Wooldridge, J. M. (2012). *Introductory Econometrics: A Modern Approach*. Ohio, USA: South-Western Cengage Learning.

Y. Wu, C. Gaunt & S. Gray. (2010). 'A Comparison of Alternative Bankruptcy Prediction Models'. *Journal of Contemporary Accounting & Economics*, *6*(1), 34-45.

Zavgren, C. V. (1985). 'Assessing the Vulnerability to Failure of American Industrial Firms: A Logistic Analysis'. *Journal of Business Finance & Accounting*, *12*(1), 19-45.

Zmijewski, M. E. (1984). 'Methodological Issues Related to the Estimation of Financial Distress Prediction Models'. *Journal of Accounting Research*, *22*, 59-82.

Databases

Bundesministerium der Justiz und für Verbraucherschutz (2017). *German Federal Gazette*. (March 2, 2017). Available at: https://www.bundesanzeiger.de/ebanzwww/wexsservlet

Bureau van Dijk (2017). *Amadeus*. (March 2, 2017). Available at: https://amadeus-bvdinfo-com.ez.hhs.se/home.serv?product=amadeusneo&loginfromcontext=ipaddress

Companies House (2017). *Companies House*. (March 2, 2017). Available at: https://www.gov.uk/government/organisations/companies-house

European Union (2015). *Statistics on European Cities*. (March 2, 2017). Avilable at: http://ec.europa.eu/eurostat/statistics-explained/index.php/File:Focus_on_European_cities_RYB2015.xlsx

FIFA Index (2017). *FIFA Index*. (March 2, 2017). Available at: https://www.fifaindex.com

Transfermarkt GmbH & Co. KG (2017). *Transfermarkt*. (March 2, 2017). Available at: http://www.transfermarkt.com/

UEFA (2017). *UEFA rankings for club competitions*. (February 8, 2017). Available at: http://www.uefa.com/memberassociations/uefarankings/country/index.html

Online Sources

Lindberg, A. (26 September 2016). FIFA 17's player ratings system blends advanced stats and subjective scouting. *ESPN FC*. Available at: http://www.espnfc.us/blog/espn-fc-united-blog/68/post/2959703/fifa-17-player-ratings-system-blends-advanced-stats-and-subjective-scouting [Last accessed: 16 April 2017]

UEFA (12 January 2017). The European Club Footballing Landscape. *UEFA*. Available at: http://www.uefa.org/MultimediaFiles/Download/OfficialDocument/uefaorg/Finance/02/42/27/91/2422791_DOWNLOAD.pdf [Last accessed: 17 April 2017]

StataCorp (2013). *mvrotate: Stata Manual Release 13*. College Station, Texas, Stata Press. Available at: http://www.stata.com/manuals13/mvrotate.pdf [Last accessed: 23.04.2017]

Appendices

Appendix 1 – Table of tested variables

Numerator	Denominator	Type	Source	Expected Sign
Spain	1	Dichotomous	Authors	+
Population (01.01.2012)	1	Ratio	Buraimo et al. (2005)	-
Division Y	1	Ordinal	Barajas & Rodríguez (2010)	+
Division Y-1	1	Ordinal	Barajas & Rodríguez (2010)	+
Division Y-2	1	Ordinal	Barajas & Rodríguez (2010)	+
Relegation Y-1	1	Dichotomous	Szymanski (2012)	+
Promotion Y-1	1	Dichotomous	Beech et al. (2010) Szymanski (2012)	Indeterminate
Short-term Movement (2 years)	1	Ratio	Szymanski (2012)	+
Short-term Movement (3 years)	1	Ratio	Szymanski (2012)	+
Long-term Movement (5 years)	1	Ratio	Szymanski (2012)	+
Star Player Y to Y-2	1	Dichotomous	Authors	-
Star Player Y	1	Dichotomous	Authors	-
Star Player Y-1	1	Dichotomous	Authors	-
Star Player Y-2	1	Dichotomous	Authors	-
Total Equity	Total Liabilities	Ratio	Altman (2000)	-
Employee Costs	Revenues	Ratio	Barajas & Rodríguez (2010) Szymanski (2012)	+
Owner's Equity	Total Assets	Ratio	Skogsvik (1987)	-
Total Liabilities	Total Assets	Ratio	Beaver 1966 Ohlson 1980 Barajas & Rodríguez 2010 Mourao 2012 Szymanski 2012	+
Employee Costs	Total Costs	Ratio	Mourao 2012	+
Amortization	Total Costs	Ratio	Mourao 2012 Ascari & Gagnepain 2006	+
Long-term Debt	Total Revenues	Ratio	Szymanski 2012	+

Numerator	Denominator	Type	Source	Expected Sign
Revenues	Total Liabilities	Ratio	Beech at al. 2010	-
Total Costs	Total Revenues	Ratio	Beech at al. 2010	+
Working Capital	Total Assets	Ratio	Altman 1968 Ohlson 1980 Altman 2000	-
Current Liabilities	Current Assets	Ratio	Barajas & Rodríguez 2010	+
Current Assets	Current Liabilities	Ratio	Marcinkowska 2013 Smijewski 1984	-
Cash Assets	Current Liabilities	Ratio	Skogsvik 1987	-
Net Income	Total Assets	Ratio	Beaver 1966 Ohlson 1980 Zmijewski 1984	-
EBIT	Total Assets	Ratio	Altman 1968 Altman 2000	-
Revenues	Total Assets	Ratio	Altman 1968 Altman 2000	-
Change in Net Income	1	Ratio	Ohlson 1980	-
INTWO	1	Dichotomous	Ohlson 1980	-
OENEG	1	Dichotomous	Ohlson 1980	Indeterminate
Total Equity	Fixed Assets	Ratio	Marcinkowska 2013	-
Change in Revenues	1	Ratio	Skogsvik 1987	-
Change in Owner's Equity	1	Ratio	Skogsvik 1987	-

Appendix 2 – Table of distressed clubs

Club	Country	Year of Distress	Month of Distress	Issue	Source
A.C.SIENA - SOCIETA' A RESPONSABILITA' LIMITATA	Italy	2014	July	Insolvency Proceedings	http://www.calciomio.fr/archive/officiel-siena-padova-font-faillite_253589.html
ALBACETE BALOMPIE, SAD	Spain	2010	April	Insolvency Proceedings	http://www.abc.es/hemeroteca/historico-06-04-2010/abc/Toledo/el-albacete-balompie-solicita-acogerse-al-concurso-de-acreedores_124680257789.html
ALEMANNIA AACHEN GMBH	Germany	2012	November	Insolvency Proceedings	https://www.welt.de/sport/fussball/article111179506/Drittligist-Alemannia-Aachen-meldet-Insolvenz-an.html
ASCOLI CALCIO 1898 - S.P.A.	Italy	2013	December	Insolvency Proceedings	http://www.repubblica.it/sport/calcio/2013/12/17/news/ascoli_fallimento-73834628/
ASSOCIAZIONE SPORTIVA BARI S.P.A.	Italy	2014	March	Insolvency Proceedings	http://www.90min.com/it/posts/1875804-caso-parma-i-principali-club-italiani-che-sono-falliti/8-bari
ASSOCIAZIONE SPORTIVA VARESE 1910 S.P.A.	Italy	2015	July	Insolvency Proceedings	http://www.football-italia.net/69209/reggina-venezia-and-varese-bankrupt
ATHLETIC CLUB ARLES AVIGNON	France	2015	July	Insolvency Proceedings	http://www.francefootball.fr/news/arles-avignon-da-pose-le-bilan/578262
C.D. TENERIFE SAD	Spain	2013	April	Liquidity Issues	http://www.laopinion.es/deportes/2013/04/06/tenerife-esquiva-concursal/468513.html
CALCIO COMO S.R.L. O CALCIO COMO 1907 S.R.L. E IN BREVE COMO CALCIO S.R.L.	Italy	2016	July	Insolvency Proceedings	http://www.italy24.ilsole24ore.com/art/panorama/2017-03-20/michael-essien-s-wife-buys-como-soccer-team-bankruptcy-proceedings-180109.php?uuid=AEhwC5p
CALCIO PORTOGRUARO-SUMMAGA S.R.L. DILETTANTISTICA IN LIQUIDAZIONE	Italy	2013	July	Insolvency Proceedings	http://www.sampnews24.com/portogruaro-ieri-la-collaborazione-con-la-samp-oggi-il-fallimento-8199/
CORDOBA CLUB DE FUTBOL SAD	Spain	2011	June	Insolvency Proceedings	http://www.diariocordoba.com/noticias/deportes/cordoba-cf-entra-oficialmente-concurso-acreedores_644634.html

Club	Country	Year of Distress	Month of Distress	Issue	Source
DSC ARMINIA BIELEFELD GMBH & CO. KGAA	Germany	2011	April	Insolvency Proceedings	http://www.spiegel.de/sport/fussball/drohende-insolvenz-arminia-bielefeld-leiht-sich-geld-vom-ligaverband-a-757090.html
DUNFERMLINE ATHLETIC FOOTBALL CLUB LIMITED	Scotland	2013	March	Insolvency Proceedings	http://www.bbc.com/sport/football/21936883
ELCHE CLUB DE FUTBOL SAD	Spain	2015	June	Insolvency Proceedings	http://www.expansion.com/valencia/2015/08/07/55c49be946163fe25b8b4587.html
FC BANÍK OSTRAVA, A.S.	Czech Republic	2013	May	Bailout	http://www.radio.cz/en/section/news/banik-ostrava-saved-from-bankruptcy
FOOTBALL CLUB ISTRES OUEST PROVENCE	France	2015	June	Insolvency Proceedings	http://www.foot-national.com/foot-dncg-un-ancien-club-pro-relegue-en-dh-72348.html
FOOTBALL PADOVA S.P.A. IN LIQUIDAZIONE	Italy	2014	July	Insolvency Proceedings	http://www.calciomio.fr/archive/officiel-siena-padova-font-faillite_253589.html
GIRONA FUTBOL CLUB, SAD	Spain	2013	July	Insolvency Proceedings	http://www.marca.com/2015/04/09/futbol/equipos/girona/1428605886.html
HEART OF MIDLOTHIAN PLC	Scotland	2013	June	Insolvency Proceedings	http://www.bbc.com/sport/football/22953448
MALAGA CLUB DE FUTBOL, SOCIEDAD ANONIMA DEPORTIVA	Spain	2012	December	Sanctions from Association	http://www.dailymail.co.uk/sport/football/article-2251715/Malaga-banned-season-European-competition-failing-pay-debts.html
MSV DUISBURG GMBH & CO. KOMMANDITGESELLSCHAFT AUF AKTIEN	Germany	2012	November	Liquidity Issues	http://www.focus.de/sport/fussball/bundesliga2/2-bundesliga-msv-duisburg-droht-die-insolvenz_aid_861661.html
N.V. ADO DEN HAAG	Netherlands	2016	January	Liquidity Issues	http://www.dutchnews.nl/news/archives/2016/04/ado-den-haag-finally-gets-its-money-from-chinese-owner/
NOTTINGHAM FOREST FOOTBALL CLUB LIMITED	England	2014	December	Sanctions from Association	https://www.theguardian.com/football/2014/dec/15/leeds-blackburn-nottingham-forest-financial-fair-play-transfer-embargo

Club	Country	Year of Distress	Month of Distress	Issue	Source
OS BELENENSES - SOCIEDADE DESPORTIVA DE FUTEBOL, S.A.D.	Portugal	2015	January	Liquidity Issues	http://www.jornaldenegocios.pt/empresas/detalhe/sad_do_belenenses_deve_74_milhoes_e_quase_um_terco_e_ao_estado
PARMA FOOTBALL CLUB SPA OPPURE IN FORMA ABBREVIATA PARMA F.C. S.P.A. IN FALLIMENTO	Italy	2015	March	Insolvency Proceedings	http://www.espnfc.com/parma/story/2500817/parma-declared-bankrupt-will-be-relegated-to-amateur-leagues https://www.theguardian.com/football/blog/2015/jun/23/parma-the-football-club-milked-for-all-their-worth
PIACENZA FOOT-BALL CLUB S.P.A.	Italy	2012	March	Insolvency Proceedings	http://www.ilpiacenza.it/cronaca/fallimento-piacenza-calcio.html
PLYMOUTH ARGYLE FOOTBALL COMPANY LIMITED(THE)	England	2011	March	Insolvency Proceedings	http://news.bbc.co.uk/sport2/hi/football/teams/p/plymouth_argyle/9414349.stm
PORTSMOUTH CITY FOOTBALL CLUB LIMITED	England	2010	February	Insolvency Proceedings	http://www.nydailynews.com/sports/more-sports/english-premier-league-portsmouth-enters-bankruptcy-sad-chapter-team-112-year-history-article-1.199756 https://www.theguardian.com/football/2010/feb/26/portsmouth-premierleague
PSV N.V.	Netherlands	2011	May	Bailout	http://www.n-tv.de/sport/fussball/Eindhoven-rettet-den-PSV-article3423636.html
REAL BETIS BALOMPIE, SOCIEDAD ANONIMA DEPORTIVA	Spain	2010	October	Insolvency Proceedings	http://www.football-espana.net/21130/lopera-took-money-out-betis-claim
REAL CLUB CELTA DE VIGO SAD	Spain	2008	November	Insolvency Proceedings	http://cincodias.com/cincodias/2008/06/11/empresas/1213191601_850215.html
REAL CLUB DEPORTIVO MALLORCA SAD	Spain	2010	June	Insolvency Proceedings	http://www.iureabogados.com/el-club-deportivo-mallorca-sale-del-concurso-de-acreedores-delta-sport-desiste-recurso.html
REAL MURCIA CF SAD	Spain	2009	February	Insolvency Proceedings	http://www.laopiniondemurcia.es/deportes/2009/02/19/declaran-concurso-voluntario-acreedores-real-murcia-demostrarse-insolvencia/151958.html

Club	Country	Year of Distress	Month of Distress	Issue	Source
REAL RACING CLUB DE SANTANDER, SAD	Spain	2011	July	Insolvency Proceedings	http://www.marca.com/2011/07/08/futbol/equipos/racing/1310124506.html
REAL SOCIEDAD DE FUTBOL SAD	Spain	2008	July	Insolvency Proceedings	http://www.diariovasco.com/20080707/deportes/beti-erreala/juzgado-declara-real-sociedad-200807071437.html
REAL VALLADOLID CLUB DE FUTBOL SA D	Spain	2011	December	Insolvency Proceedings	http://www.sportsbusinessdaily.com/Global/Issues/2014/07/24/Franchises/Notes.aspx?printandclose=true
REAL ZARAGOZA SAD	Spain	2011	June	Insolvency Proceedings	http://www.cbc.ca/sports/soccer/spanish-club-real-zaragoza-files-for-bankruptcy-protection-1.1116328 http://www.football-espana.net/46166/zaragoza-facing-bankruptcy
REGGINA CALCIO - S.P.A. -	Italy	2015	July	Insolvency Proceedings	http://www.football-italia.net/69209/reggina-venezia-and-varese-bankrupt
RFC 2012 P.L.C.	Scotland	2012	February	Insolvency Proceedings	http://www.ibtimes.co.uk/glasgow-rangers-bankrupt-protection-administration-football-299005
RTS WIDZEW ŁÓDŹ S.A. W UPADŁOŚCI LIKWIDACYJNEJ	Poland	2015	June	Insolvency Proceedings	http://www.polskieradio.pl/43/265/Artykul/1475022,Widzew-Lodz-reaktywacja-Zasluzony-klub-zacznie-od-czwartej-ligi
RUCH CHORZÓW S.A.	Poland	2016	December	Liquidity Issues	http://www.tylkoekstraklasa.pl/ruch-chorzow-bardzo-blisko-bankructwa/
S.P.A.L. 2013 S.R.L.	Italy	2012	July	Insolvency Proceedings	http://www.football-italia.net/83269/spal-promoted-serie-b
SHEFFIELD WEDNESDAY FOOTBALL CLUB LIMITED	England	2010	November	Bailout	http://www.telegraph.co.uk/sport/football/teams/sheffield-wednesday/8168128/Sheffield-Wednesday-avoid-administration-after-Leicester-City-chairman-Milan-Mandaric-completes-takeover.html
SOUTHAMPTON FOOTBALL CLUB LIMITED	England	2009	April	Insolvency Proceedings	https://www.theguardian.com/football/2009/apr/02/southampton-leisure-holdings-administration
SOUTHEND UNITED FOOTBALL CLUB LIMITED(THE)	England	2009	October	Liquidity Issues	https://www.theguardian.com/football/2009/oct/27/southend-hmrc-administration-hearing

Club	Country	Year of Distress	Month of Distress	Issue	Source
SPORTING CLUBE DE PORTUGAL - FUTEBOL, SAD	Portugal	2016	September	Liquidity Issues	http://www.jn.pt/desporto/interior/sporting-precisa-de-44-milhoes-para-evitar-falencia-5417939.html
STICHTING FC TWENTE '65	Netherlands	2014	April	Liquidity Issues	http://www.football-oranje.com/fc-twente-dire-financial-straits/
THE BLACKBURN ROVERS FOOTBALL AND ATHLETIC LIMITED	England	2014	December	Sanctions from Association	http://www.bbc.com/sport/football/34996365
THE DUNDEE FOOTBALL CLUB LIMITED	Scotland	2010	October	Insolvency Proceedings	https://www.theguardian.com/football/blog/2010/oct/15/dundee-future-administration
UNIÃO DESPORTIVA DE LEIRIA FUTEBOL, SAD	Portugal	2011	December	Liquidity Issues	http://relvado.aeiou.pt/1-liga/joao-bartolomeu-se-nao-conseguirmos-pagar-aos-jogadores-paciencia-302231
WISŁA KRAKÓW S.A.	Poland	2012	October	Liquidity Issues	http://sportowefakty.wp.pl/pilka-nozna/319114/wisla-krakow-na-skraju-bankructwa
WROCŁAWSKI KLUB SPORTOWY ŚLĄSK WROCŁAW S.A.	Poland	2013	September	Insolvency Proceedings	http://www.tvn24.pl/wroclaw,44/to-droga-do-bankructwa-i-likwidacji-slaska-wroclaw,355334.html

Appendix 3 – Two-year model: descriptive statistics for variables

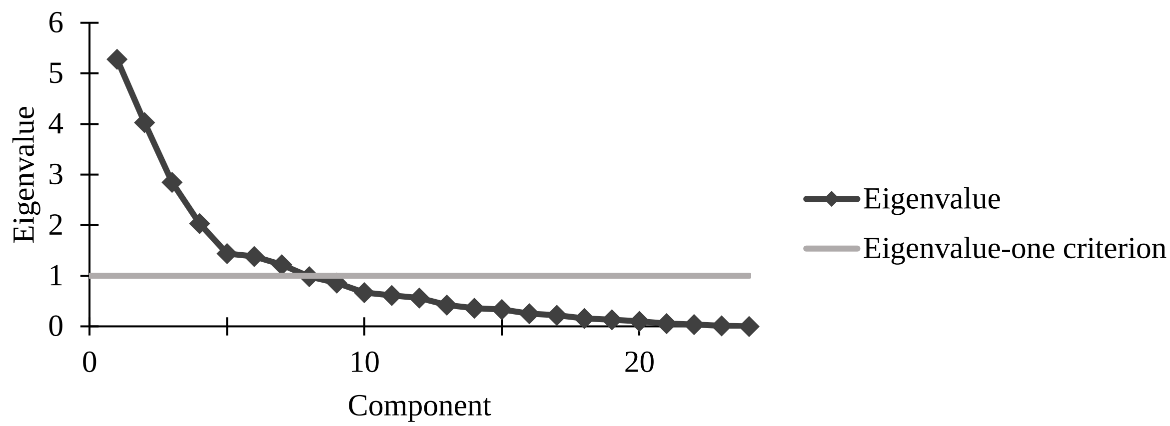
	Mean		Median		Standard Deviation	
Variable	**Survival**	**Distress**	**Survival**	**Distressed**	**Survival**	**Distress**
Spain	0.0816	0.2857	0.0000	0.0000	0.2752	0.4564
Population (01.01.2012)	755,929	361,278	175,119	257,300	1,772,582	369,978
Division Y-1	1.5612	1.6327	1.0000	2.0000	0.6745	0.6355
Division Y-2	1.5816	1.5102	1.0000	1.0000	0.6723	0.5818
Division Y-3	1.6122	1.6327	1.0000	2.0000	0.7270	0.7273
Relegation Y-2	0.1020	0.1224	0.0000	0.0000	0.3043	0.3312
Promotion Y-2	0.1224	0.0000	0.0000	0.0000	0.3295	0.0000
Short-term Movement (2 years)	0.3980	0.2857	0.0000	0.0000	0.6380	0.5774
Short-term Movement (3 years)	0.6122	0.5918	0.0000	0.0000	0.8076	0.7615
Long-term Movement (4 years)	0.7959	0.8367	0.0000	1.0000	0.9944	0.8743
Star Player Y-1 to Y-3	0.1531	0.0408	0.0000	0.0000	0.3619	0.1999
Star Player Y-1	0.0714	0.0000	0.0000	0.0000	0.2589	0.0000
Star Player Y-2	0.0918	0.0204	0.0000	0.0000	0.2903	0.1429
Star Player Y-3	0.1224	0.0408	0.1885	0.0000	0.3295	0.1999
BV Equity / BV Liabilities	0.5163	-0.0166	0.6247	0.0220	1.4330	0.3896
Wages / Revenues	0.6645	0.7194	0.1586	0.7324	0.2754	0.2464
Owner's Equity / Assets	0.0132	-0.2969	0.8414	0.0215	0.8934	0.8932
Total Liabilities / Total Assets	0.9868	1.2968	0.5667	0.9785	0.8934	0.8931
Wages / Total Costs	0.5419	0.5790	0.1009	0.5704	0.1471	0.1504
Amortization / Total Costs	0.1164	0.1317	1.2833	0.1043	0.0725	0.0918
Revenues / Total Liabilities	1.6603	1.0439	1.1270	0.7498	1.3727	1.1054
Total Costs / Revenues	1.2306	1.2521	0.0216	1.2449	0.3794	0.3534
Working capital / Total assets	0.0518	0.0471	1.2557	0.0334	0.1503	0.1560
Current Liabilities / Current Assets	2.3978	3.4637	0.7967	1.6269	3.7838	4.6305
Current Assets / Current Liabilities	1.0034	0.5786	0.0466	0.6147	0.8816	0.3402
Cash assets / Current Liabilities	0.2629	0.0904	-0.0334	0.0169	0.5437	0.1455
Net Income / Total Assets	-0.1109	-0.1512	-0.1143	-0.0410	0.2881	0.3171
EBIT / Total Assets	-0.1844	-0.2033	1.0209	-0.1450	0.3624	0.3423
Revenues / Total Assets	1.2252	1.0587	0.3491	0.8199	0.8785	0.9220
Equity / Fixed Assets	0.5295	-0.4242	0.0298	0.0534	2.4515	1.3776
Long-term Debt / Revenues	0.3707	0.4415	0.0000	0.1217	0.9995	0.8198

Appendix 4 – Two-year model: correlations between all variables

	Spain	Population	Division (Y-1)	Division (Y-2)	Division (Y-3)	Relegation (Y-2)	Promotion (Y-2)	Short-term Movement (2 years)	Short-term Movement (3 years)	Long-term Movement (4 years)	Star Player (Y-1 to Y-3)	Star Player (Y-1)	Star Player (Y-2)	Star Player (Y-3)	BV Equity / BV Liabilities	Wages / Revenues
Spain	**1**															
Population	-0.0526	**1**														
Division (Y-1)	-0.0586	-0.1771	**1**													
Division (Y-2)	-0.0379	-0.2415	**0.7997**	**1**												
Division (Y-3)	0.0276	-0.2422	**0.7470**	**0.8260**	**1**											
Relegation (Y-2)	-0.0154	0.0567	0.3461	-0.1533	-0.0147	**1**										
Promotion (Y-2)	0.0381	-0.0837	-0.1611	0.2710	0.1328	-0.0911	**1**									
Short-term Movement (2 years)	0.0688	-0.1035	0.3235	0.2936	0.4010	**0.5369**	**0.5475**	**1**								
Short-term Movement (3 years)	0.1169	-0.0980	0.3403	0.3358	0.4309	0.4283	0.4804	**0.8402**	**1**							
Long-term Movement (4 years)	0.1114	-0.1030	0.4157	0.4039	0.4838	0.3840	0.4036	**0.7485**	**0.9087**	**1**						
Star Player (Y-1 to Y-3)	0.0272	0.3553	-0.2932	-0.3151	-0.3060	-0.0516	-0.0977	-0.1734	-0.2477	-0.2836	**1**					
Star Player (Y-1)	-0.0043	0.1534	-0.2013	-0.1948	-0.1892	-0.0754	-0.0604	-0.1286	-0.1698	-0.1893	**0.6183**	**1**				
Star Player (Y-2)	0.0381	0.2736	-0.2022	-0.2354	-0.2286	-0.0018	-0.0730	-0.1115	-0.1709	-0.2000	**0.7471**	**0.5739**	**1**			
Star Player (Y-3)	0.0588	0.3884	-0.2569	-0.2827	-0.2746	-0.0328	-0.0877	-0.1489	-0.2170	-0.2501	**0.8972**	**0.5804**	**0.6486**	**1**		
BV Equity / BV Liabilities	-0.1023	-0.1377	-0.1342	-0.1117	-0.1107	-0.0846	-0.0352	-0.0852	-0.1176	-0.1389	-0.0243	0.0313	-0.0294	-0.0195	**1**	
Wages / Revenues	0.1354	0.0258	0.2107	0.2366	0.1749	-0.0859	-0.0517	-0.0828	0.0160	0.0107	-0.1073	-0.1166	-0.0822	-0.0823	-0.195	**1**
Owner's Equity / Assets	-0.0264	-0.3568	-0.0286	-0.0038	-0.0079	-0.0013	0.0633	0.0608	0.0357	0.0088	-0.0647	0.0965	0.0627	-0.0894	0.4193	-0.2952
Total Liabilities / Total Assets	0.0265	0.3568	0.0286	0.0039	0.0079	0.0013	-0.0633	-0.0608	-0.0357	-0.0089	0.0648	-0.0965	-0.0627	0.0894	-0.4193	0.2951
Wages / Total Costs	0.3228	0.0389	0.0861	0.0389	0.1130	0.0604	-0.0531	0.0614	0.1317	0.1156	-0.0389	-0.0591	-0.0288	-0.0327	-0.167	**0.5822**
Amortization / Total Costs	0.1797	0.2130	-0.2707	-0.2378	-0.2439	-0.0083	0.0920	-0.0089	-0.0421	-0.0800	0.3065	0.1524	0.2707	0.3278	-0.1042	0.0939
Revenues / Total Liabilities	-0.2515	-0.1424	0.1199	0.0679	0.0729	0.0488	-0.0817	-0.0511	-0.0179	0.0233	-0.0355	0.0108	-0.0175	-0.1018	0.3749	-0.2739
Total Costs / Revenues	-0.0606	-0.0033	0.1757	0.2543	0.1196	-0.1533	0.0036	-0.1162	-0.0533	-0.0631	-0.1153	-0.1005	-0.0777	-0.0858	-0.0759	**0.7625**
Working capital / Total assets	0.2654	-0.0472	-0.0626	-0.0957	-0.0924	0.0545	-0.0143	0.0004	0.0058	-0.0030	0.0571	0.1005	-0.0056	0.0908	0.0394	-0.1596
Current Liabilities / Current Assets	-0.0190	0.0668	0.1110	0.1196	0.0618	-0.0871	-0.0906	-0.0750	-0.0982	-0.0394	-0.0884	-0.0965	-0.0826	-0.0648	-0.2267	0.3050
Current Assets / Current Liabilities	-0.1245	-0.1128	-0.0632	-0.0438	-0.0582	-0.0380	0.0078	-0.0610	-0.0815	-0.1041	0.0776	0.0624	0.1072	-0.0113	0.3933	-0.2799
Cash assets / Current Liabilities	-0.0359	-0.0414	-0.0696	-0.0775	-0.0631	-0.0504	-0.0761	-0.0804	-0.0757	-0.0977	0.0211	0.0423	0.0271	-0.0192	0.3401	-0.0645
Net Income / Total Assets	0.1339	-0.0582	-0.1805	-0.2637	-0.1604	0.1675	0.0018	0.1133	0.0505	0.0099	0.1384	0.0935	0.1598	0.0615	0.0974	**-0.5193**
EBIT / Total Assets	0.1504	-0.0191	-0.1951	-0.2637	-0.0993	0.1203	-0.0166	0.1057	0.0430	-0.0104	0.1205	0.1020	0.1259	0.0616	0.0942	**-0.5832**
Revenues / Total Assets	-0.2135	-0.0272	0.1887	0.0855	0.1067	0.1642	-0.0777	0.0255	0.0750	0.1308	-0.0043	-0.0383	0.0077	-0.0957	-0.1018	-0.1879
Equity / Fixed Assets	-0.0803	-0.2311	0.0089	0.0389	0.0196	-0.0530	0.0118	-0.0171	-0.0567	-0.0781	-0.0610	0.0428	0.0079	-0.0691	0.3896	-0.2440
Long-term Debt / Revenues	0.2238	0.1309	0.0636	0.0609	0.1022	-0.0633	-0.0874	-0.0713	-0.1135	-0.0517	-0.0031	-0.0737	-0.0747	0.0205	-0.2229	0.2585

	Owner's Equity / Assets	Total Liabilities / Total Assets	Wages / Total Costs	Amortization / Total Costs	Revenues / Total Liabilities	Total Costs / Revenues	Working capital / Total assets	Current Liabilities / Current Assets	Current Assets / Current Liabilities	Cash assets / Current Liabilities	Net Income / Total Assets	EBIT / Total Assets	Revenues / Total Assets	Equity / Fixed Assets	Long-term Debt / Revenues
Owner's Equity / Assets	**1**														
Total Liabilities / Total Assets	**-1.0000**	**1**													
Wages / Total Costs	-0.1822	0.1822	**1**												
Amortization / Total Costs	0.0334	-0.0334	0.1735	**1**											
Revenues / Total Liabilities	0.3481	-0.3480	-0.0999	-0.4352	**1**										
Total Costs / Revenues	-0.2070	0.2069	-0.0522	-0.0020	-0.2746	**1**									
Working capital / Total assets	0.1417	-0.1416	-0.0388	0.0471	-0.0596	-0.1536	**1**								
Current Liabilities / Current Assets	-0.4252	0.4252	0.1210	0.0096	-0.3587	0.2670	-0.2039	**1**							
Current Assets / Current Liabilities	0.3497	-0.3497	-0.2707	-0.1775	**0.5191**	-0.1108	0.2322	-0.4342	**1**						
Cash assets / Current Liabilities	0.2151	-0.2151	0.0075	-0.1210	0.3427	-0.0918	0.0298	-0.2075	**0.7058**	**1**					
Net Income / Total Assets	0.3064	-0.3063	0.0202	0.1533	0.0714	**-0.6597**	0.2877	-0.1509	0.1842	0.1023	**1**				
EBIT / Total Assets	0.2037	-0.2036	0.0089	0.1583	0.0006	**-0.7266**	0.2456	-0.0728	0.1210	0.0586	**0.8458**	**1**			
Revenues / Total Assets	-0.0073	0.0073	0.0123	-0.4603	**0.7989**	-0.2700	-0.0911	-0.2629	0.2822	0.1131	0.0194	-0.0813	**1**		
Equity / Fixed Assets	**0.6562**	**-0.6561**	-0.1659	-0.0547	0.4919	-0.1474	0.1043	-0.3105	**0.5307**	0.3231	0.1824	0.1756	0.1053	**1**	
Long-term Debt / Revenues	**-0.5490**	**0.5491**	0.1816	0.0587	-0.3306	0.1619	0.0024	0.3611	-0.2191	-0.1337	-0.0802	-0.0440	-0.2385	-0.3358	**1**

Appendix 5 – Two-year model: components and eigenvalue-one criterion



Appendix 6 – Two-year model: coefficients of variables above 0.3 for the components with eigenvalues above 1 after the promax rotation

	Component						
Variable	**1**	**2**	**3**	**4**	**5**	**6**	**7**
Population (01.01.2012)		-0.3026					
Division Y-1		0.4956					
Division Y-2		0.5194					
Division Y-3		0.5051					
Short-term Movement (2 years)				0.5393			
Short-term Movement (3 years)				0.5817			
Long-term Movement (4 years)				0.5343			
BV Equity / BV Liabilities							
Wages / Revenues			-0.3281				0.476
Owner's Equity / Assets	-0.5361						
Total Liabilities / Total Assets	0.5362						
Wages / Total Costs							0.8105
Amortization / Total Costs					-0.3454		
Revenues / Total Liabilities					0.5213		
Total Costs / Revenues			-0.4983				
Working capital / Total assets							
Current Liabilities / Current Assets							
Current Assets / Current Liabilities						0.5909	
Cash assets / Current Liabilities						0.6218	
Net Income / Total Assets			0.5207				
EBIT / Total Assets			0.5536				
Revenues / Total Assets					0.6521		
Equity / Fixed Assets	-0.3272						
Long-term Debt / Revenues	0.3824						

Appendix 7 – Two-year model: component interpretations

Component	Content
1	Capital structure
2	Division
3	Profitability
4	Movements (relegations and promotions) in the last years
5	Return on assets
6	Current liabilities coverage
7	Spending on players