The Impact of Player Turnover on Football Teams Performance

Pedro Miguel Branco Pires

Thesis to obtain the Master of Science Degree in

Industrial Engineering and Management

Supervisor: Prof. António Sérgio Ribeiro

Examination Committee

Chairperson: Prof. Miguel Simões Torres Preto

Supervisor: Prof. António Sérgio Constantino Folgado Ribeiro

Member of the Committee: Prof. Hugo Miguel Fragoso de Castro Silva

September 2021

Declaration

I declare that this document is an original work of my authorship and that it fulfils all the requirements of the Code of Conduct and Good Practices of the Universidade de Lisboa.

Declaração

Declaro que o presente documento é um trabalho original da minha autoria e que cumpre todos os requisitos do Código de Conduta e Boas Práticas da Universidade de Lisboa.

Abstract

In the past decades, massive transformations of football clubs into authentic public companies have been witnessed. This included the critical need of an excellent financial and asset management, as well as players becoming the most crucial club asset. Moreover, across the past decade, it became apparent that the capacity of clubs to conduct player transfers is rising. On the other hand, a problem arises when a club undergoes several changes from one season to another: team performance drops.

This masters' dissertation addresses team performance and club stability by applying Data Envelopment Analysis and Linear Regression Model. The main goal is to update the literature regarding the topic, and eventually attempt to help football clubs make decisions. Both models have as the primary input variable the percentage of new players minutes and as the output variable the variation of points regarding the previous season.

Concerning the results, it is important to highlight that the team with the lowest %NPMin and the team with the highest variation of points regarding the previous season are permanently assigned maximum efficiency. Additionally, 58% of the cases (average from all leagues) with less than 20%NPMin had efficiency equal to 1. From the results of DEA models, no direct conclusions could be drawn, only an evident tendency line reflecting the relation between %NPMin and VarPts. Ultimately, the final linear regression model allowed to prove that stability affects the team's performance: clubs that have fewer new players fielded are more likely to have a better performance than before.

Keywords: Player turnover; Team Stability; Football; Performance; DEA; Regression Model

Resumo

Nas últimas décadas, tem se assistido a uma gradual transformação dos clubes de futebol em empresas autênticas, onde uma boa gestão financeira e patrimonial é crucial. Por sua vez, os jogadores tornaram-se o ativo mais valioso dos clubes. Além disso, com o evoluir do século XXI, surge um problema relacionado com o desempenho da equipa quando um clube tem muitas mudanças no seu plantel de uma época para a seguinte.

Esta dissertação aborda o desempenho e estabilidade do clube através da aplicação de *Data Envelopment Analysis* e do Modelo de Regressão Linear. O objetivo principal passa por atualizar a literatura relativa ao tema e eventualmente ajudar os clubes de futebol a tomarem melhores decisões. Ambos os modelos têm como principal variável a %NPMin e como variável de ouput VarPts em relação à época anterior.

Os principais destaques nos resultados são: a equipa com a %NPMin mais baixa e a equipa com a maior VarPts em relação à época anterior recebem sempre eficiência máxima. Além disso, 58% dos casos (média de todas as ligas) com menos de 20% de NPMin tiveram uma eficiência igual a 1. A partir dos modelos DEA, apenas foi possível ver uma linha de tendência evidente que refletia a relação entre %NPMin e VarPts. Finalmente, com o modelo final da regressão linear foi possível provar que a estabilidade afeta o desempenho da equipa. Os clubes que têm menos jogadores novos em campo têm mais probabilidades de ter um melhor desempenho do que antes.

Keywords: Rotatividade dos jogadores; Estabilidade da equipa; Futebol; Desempenho; DEA; Modelo de Regressão

Acknowledgements

I would like to express my gratitude to the persons listed below, without whom I would not have been able to conduct this study as well as complete my master's degree.

Firstly, I would like to thank my supervisor, Professor António Sérgio Ribeiro for his guidance, insights and all the provided knowledge and feedback throughout the study. All your interesting methods and tools, and your stimulating questions have made this masters' dissertation complete.

I would also like to thank my parents and family who have been consistently supportive for the previous year and who had to put up with my worries and stress.

A special thanks to my sister Joana who has always been an example to follow in all respects. Our share of ideas and discussions inspired me to think about the bigger picture and from many viewpoints in order to develop a complete and impartial analysis.

Furthermore, I would also want to convey my appreciation to all my friends and colleagues from Instituto Superior Técnico who have made this wonderful experience more pleasant and fun. Thank you for becoming my second family and growing alongside me, your positive attitude and genuine friendship are the most memorable aspects of this journey.

Table of Contents

ABSTRACT III

RESUMO IV

ACKNOWLEDGEMENTS V

TABLE OF CONTENTS VI

LIST OF FIGURES VIII

LIST OF TABLES IX

LIST OF ACRONYMS X

1. INTRODUCTION 1

1.1 Problem Contextualization 1
1.2 Dissertation Goals 1
1.3 Dissertation Structure 2

2. FOOTBALL INDUSTRY 4

2.1 Popularity 4
2.2 Players contracts and transfers 5
2.3 Football Finances 7
2.3 "It is all about performance" mentality 9
2.4 Lack of studies regarding player turnover vs performance 10
2.5 Football players demographs 11

3. LITERATURE REVIEW 14

3.1 Job turnover in general companies 14
3.2 Manager turnover in sports 18
3.3 Squad stability, player turnover and team performance 23
3.4 Productivity, Effectiveness and Efficiency 26
3.5 Chapter conclusions 30

4. METHODOLOGY 31

4.1 PERFORMANCE EVALUATION METHODS **31**
4.1.1 DATA ENVELOPMENT ANALYSIS – DEA 32
4.1.1.1 DEA – CRS Model 33
4.1.1.2 DEA – VRS Model 36
4.1.2 MULTIPLE LINEAR REGRESSION 37
4.1.2.1 Least-Square Method 38
4.1.2.2 Model's significance (F-test) 39
4.1.2.3 Coefficient of Determination (R2) 40
4.1.2.4 Variable's significance 41
4.1.2.5 Model with Interaction Terms 41

5. METHODOLOGY APPLICATION 43

5.1 DATA COLLECTION **43**
5.2 VARIABLES **44**
5.3 MODEL **49**

6. RESULTS 51

6.1 DEA RESULTS **51**
6.2 LINEAR REGRESSION RESULTS **60**
6.3 CHAPTER CONCLUSIONS **63**

7. CONCLUSIONS AND FUTURE CONSIDERATIONS 64

7.1 LIMITATIONS AND DIFFICULTIES **65**
7.2 FUTURE CONSIDERATIONS **65**

REFERENCES 66

APPENDIX A – CORRELATION TABLE 74

APPENDIX B – DEA RESULTS FOR 2017-2018 75

APPENDIX C – DEA RESULTS FOR 2016-2017 76

List of Figures

FIGURE 1 – MAP OF SPORTS' POPULARITY (ADAPTED FROM A 2006 NATIONAL GEOGRAPH GRAPH CALLED "SOCCER UNITED THE WORLD.") 4

FIGURE 2 - NUMBER OF TRANSFERS PER YEAR ACHIEVING THE ALL-TIME TOP 100 (ADAPTED FROM BBC.COM) 8

FIGURE 3 - HIGHEST PAID FOOTBALL PLAYERS IN 2020 (ADAPTED FROM FORBES) 9

FIGURE 4 - % OF PLAYERS JOINED THE FIRST TEAM AFTER JANUARY 1ST, 2016. (CIES, 2016B) 11

FIGURE 5 - AVERAGE SIGNINGS IN 2016 BY LEAGUE (CIES, 2016A) 12

FIGURE 6 - COMPARISON BETWEEN THE YEAR OF RECRUITMENT OF PLAYERS IN 2009 AND 2019. ADAPTED FROM CIES MONTHLY REPORT 49 13

FIGURE 7 - AVERAGE % NEW SIGNINGS FOR CHAMPIONS ON SIGNIFICANT FOOTBALL LEAGUES (2009-2017) 24

FIGURE 8 - THE EFFECT OF DIFFERENT LEVELS OF EFFICIENCY AND EFFECTIVENESS ADAPTED FROM MOUZAS (2006) 27

FIGURE 9 - PRODUCTION PROCESS AND DU 28

FIGURE 10 - VARIABLE AND CONSTANT INCOMES EXAMPLE (COELLI, 1996) 29

FIGURE 11 - TECHNICAL AND ALLOCATIVE EFFICIENCY (CARVALHO & MARQUES, 2007) 29

FIGURE 12 - PURE TECHNICAL EFFICIENCY AND SCALE EFFICIENCY (CARVALHO & MARQUES, 2007) . 30

FIGURE 13 - PERFORMANCE EVALUATION METHODS ADAPTED FROM (R. C. MARQUES, 2011) 31

FIGURE 14 - DEA-CCR MODEL EXAMPLE OF PRODUCTION FRONTIER (KIM & HARRIS, 2008) 34

FIGURE 15 - DEA FORMULATION FOR CRS METHODOLOGY (COOPER ET AL., 2006) 34

FIGURE 16 - DEA PRIMAL MODEL FORMULATION FOR CRS METHODOLOGY (COOPER ET AL., 2006) . 35

FIGURE 17 - DEA DUAL MODEL FORMULATION FOR CRS METHODOLOGY (COOPER ET AL., 2006) 35

FIGURE 18 - DEA-BCC MODEL EXAMPLE OF PRODUCTION FRONTIER (KIM & HARRIS, 2008) 36

FIGURE 19 - DEA PRIMAL MODEL FORMULATION FOR VRS METHODOLOGY (COOPER ET AL., 2006).. 37

FIGURE 20 - DEA DUAL MODEL FORMULATION FOR VRS METHODOLOGY (COOPER ET AL., 2006)..... 37

FIGURE 21 - DEA-VRS STEPS 50

FIGURE 22 - TENDENCY LINE ON %NPMIN VS EFFICIENCY 2019 54

FIGURE 23 - TENDENCY LINE ON %NPMIN VS EFFICIENCY 2018 55

FIGURE 24 - TENDENCY LINE ON %NPMIN VS EFFICIENCY 2017 55

List of Tables

TABLE 1 - STATISTICS FROM USED DATASET 47
TABLE 2 - DEA-VRS RESULTS OF EPL 2019 WITH %NPMIN AND VARPTS 52
TABLE 3 - DEA-VRS RESULTS OF EPL 2019 WITH %NPMIN (WITHOUT LIVERPOOL) 53
TABLE 4 - DEA-VRS RESULTS FOR ALL LEAGUES IN 2019 54
TABLE 5 - DEA-VRS RESULTS OF BUNDESLIGA 2019 WITH ALL VARIABLES 56
TABLE 6 - DEA-VRS RESULTS OF BUNDESLIGA 2019 WITH %NPMIN AND SQVALUE 57
TABLE 7 - DEA-VRS RESULTS OF BUNDESLIGA 2019 WITH %NPMIN AND PFIELDED 58
TABLE 8 - SUM OF VARIATION OF POINTS WITH %NPMIN BELOW 20% 58
TABLE 9 - AVERAGE OF VARIATION OF POINTS WITH %NPMIN BELOW 20% 59
TABLE 10 - LINEAR REGRESSION MODELS WITH ONE INPUT VARIABLE 60
TABLE 11 - LINEAR REGRESSION MODELS WITH TWO INPUT VARIABLES 61
TABLE 12 - LINEAR REGRESSION MODELS WITH THREE OR MORE INPUT VARIABLES 62

List of Acronyms

ANOVA	Analysis of Variance
BCC (DEA)	Banker, Charnes and Cooper DEA model
CCR (DEA)	Charnes, Cooper and Rhodes DEA model
CEO	Chief Executive Officer
CIES	International Centre for Sports Studies
COLS	Corrected Ordinary Least Squares
CRS	Constant Returns To Scale DEA model
DEA	Data Envelopment Analysis
dF	Degrees of Freedom
DFA	detrended fluctuation analysis
DMU	Decision-Making Unit
DU	Decision Unit
EPL	English Premier League
FDH	Free Disposal Hull
FIFA	Federation Internationale de Football Association
HR	Human Resources
ITC	International Transfer Certificate
MLR	Multiple Linear Regression
NBA	National Basketball Association
NCAA	National Collegiate Athletic Association
NFL	National Football League
OLS	Ordinary Least Squares

R&D	Research and Development
SE	Scale Efficiency
SFA	Stochastic Frontier Analysis
SS_E	Sum of Squares
TE	Technical Efficiency
UEFA	Union of European Football Associations
VRS	Variable Returns To Scale

1. Introduction

This chapter aims to introduce this masters' dissertation topic related to investigating player turnover on football teams. The chapter is divided into three sections: section 1.1 aims to introduce the problem contextualization, then section 1.2 provides some objectives this dissertation hopes to achieve. Finally, section 1.3 exposes the structure adopted in the following chapters of the dissertation.

1.1 Problem Contextualization

Recently, more concrete on the past decades, people have witnessed a massive transformation of football clubs into authentic public companies, where good financial and asset management are crucial. The revenues related to the clubs (both broadcast and sponsorships) are increasing significantly. The players are worth a lot more, and the prizes and sponsorships are getting bigger and bigger as time passes. With all this evolution, the managers and directors of the clubs want the club to be as efficient as possible both financially and on the field. The squad's management in terms of transfers – both sales and acquisitions – must be very well thought out towards generating full sporting and financial return from the players to achieve that desired efficiency.

Generally, small clubs invest in young and less costly talent, shape and evolve the player until some bigger club makes an irrefutable offer. On the other hand, bigger clubs look for more mature and capable players that instantly allow better services. Consequently, these smaller clubs are like a showcase of players for the "big sharks" – a common term used to describe the most prominent and wealthiest clubs. On observing this high rotation of players through teams, a problem arises - the turnover impact on performance. That impact can be on the group's chemistry, can be on general performance or even financially. Thus, it becomes crucial to analyse how turnover can disrupt a team and how a team can efficiently deal with an inevitable situation.

1.2 Dissertation Goals

The contextualization exposed on the previous topic is the motivation for the present study, which objective is to apply Data Envelopment Analysis and Linear Regression to analyse the player turnover on football teams. By applying these models, it is expected to conclude, support or corroborate the overall conclusions presented in the literature of the theme. This dissertation aims to cover these topics:

- ✓ Characterize the industry of football worldwide and focus on European leagues. The study will cover the six best leagues of the world – Spain, England, Germany, Italy, France and Portugal – with data from 2015 to 2019.

- ✓ Review previous research regarding turnover, focused more on sports, which is not so explored by the literature. This literature review aims to scrutinise the general insights given regarding models, variables, and data to analyse player turnover and team performance.
- ✓ Apply Data Envelopment Analysis to conclude regarding the teams' efficiency ratings.
- ✓ Apply Linear Regression Models to conclude regarding the relationship between input and outputs. In addition, models are expected to prove and add robustness to the DEA results
- ✓ Give insights regarding possible future work related to the same topic

1.3 Dissertation Structure

This dissertation is structured in 7 chapters:

1. **Chapter 1 – Introduction**

   The purpose of the present chapter is to introduce, contextualize and motivates the problem this dissertation aims to analyse deeply. Finally, it also shows its main goals and expected outcomes.

2. **Chapter 2 – Football Industry**

   This chapter goes more deeply into the industry of football and understands the greatness of the topic of the project. It also shows why analysing turnover vs performance on football teams can be beneficial to study. Finally, a detailed explanation of football finances and players contracts is shown.

3. **Chapter 3 – Literature Review**

   This chapter aims to survey the past studies and literature regarding productivity, effectiveness, efficiency, job turnover, manager turnover, squad stability and turnover and, finally, collect insights regarding models and variables to apply the intended methods better.

4. **Chapter 4 – Methodology**

   The methodology chapter describes the performance evaluation methods theoretically. This includes the data envelopment analysis (DEA) tool and the multiple linear regression tool. When describing the DEA tool, both the CRS and VRS models are studied in detail. Regarding the multiple linear regression tool, five models/definitions were explained: the least-square method, the model's significance (F-test), the coefficient of determination (R2), the variable's significance, and lastly, the model with interaction terms.

5. Chapter 5 – Methodology Application

The methodology application chapter thoroughly explains and demonstrates the appropriate methodology and the main variables used in the study. Accordingly, the data sets choosing process, how they were used, and how they were collected is outlined. Followed by the explanation of the choosing variables process and the execution of the model.

6. Chapter 6 – Results

In this chapter, the results of the present study are displayed. This includes the DEA results and the linear regression results. Furthermore, there is also a short conclusion regarding these topics.

7. Chapter 7 – Conclusions and Future Considerations

This last chapter ultimately presents the complete analysis of the study summarized and the primary and final views constructed across the masters' dissertation. Here the main limitations are introduced, and the future considerations are elaborated.

2. Football Industry

The following chapter will justify the motivation for the present master's dissertation and further contextualize the football environment and the problem the study seeks to answer.

2.1 Popularity

According to all studies, it is unanimous that football is the most popular sport in the world. The number of followers differs slightly from study to study. However, it is believed to be around 4 billion followers worldwide. Many factors make football the most popular and viewed sport globally. Based on SportsShow (2020) and TotalSportek (2017), some will be shown further. As mentioned before, football has the most significant fan base and audience across the world.

Similarly, it is always possible to watch football on TV, where the biggest broadcast deals and sponsorship deals regarding sports happen. The popularity and presence on the Internet and in the media are also quite evident. Furthermore, football is the most popular sport in a vast number of countries, as can be observed in figure 1, where countries with football as their most popular sport are presented with the colour green. Moreover, it is the most played sport since it has the most significant professional and amateur leagues. This number of leagues can be explained due to the accessibility to play the sport since it does not require huge investments. In contrast, the salary of the football players makes them within the highest salaries among athletes. Finally, the most followed athlete is a football player – Cristiano Ronaldo, with more than 300 million followers on Instagram.

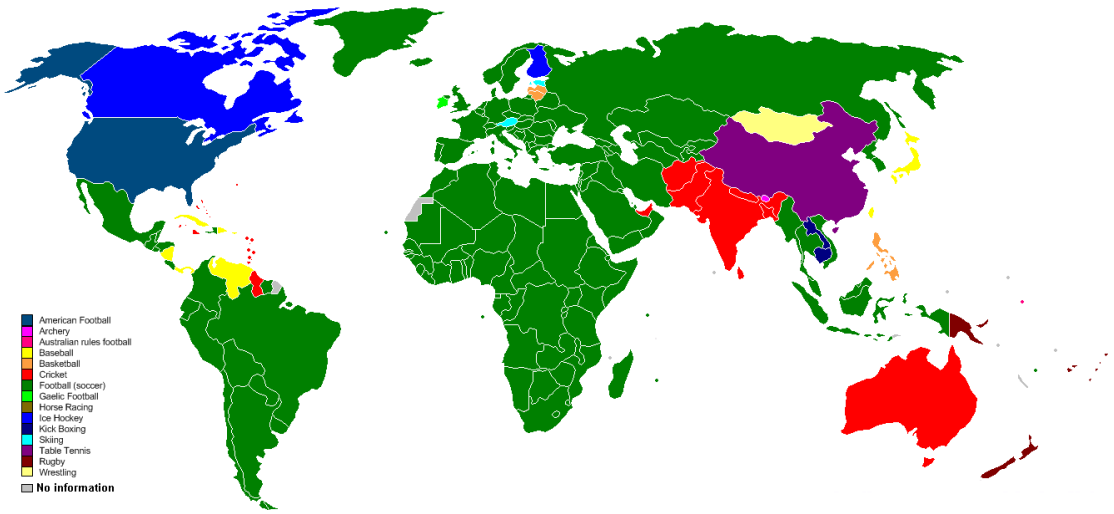

Figure 1 – Map of sports' popularity (Adapted from a 2006 National Geograph graph called "Soccer United the World.")

2.2 Players contracts and transfers

The information regarding players contracts and their specifications, incentivizing players to sign, is somehow difficult to find. However, it is important to highlight what a football player contract includes and explain all the processes of transferring players from one club to another. Additionally, the modern football world is facing a rise of complex decision-making mechanisms in the transfer market. On the one hand, clubs have to compete amongst themselves to recruit the most promising players strongly. On the other hand, they also need to consider the rise in professional football players' bargaining power.

Moreover, the stakeholders involved in these negotiations adopt fewer formal ways to negotiate and sign contracts, e.g., via electronic devices, such as emails, WhatsApp, and others. At the same time, this can speed up the terms of negotiation. In contrast, it can also create some disorder, mainly due to the number of bureaucratic documents sent and resented that can get lost. It is acknowledged that clubs discuss multiple player options and negotiate many possible signings simultaneously. However, players and agents do so as well. Thus, this previously described disorder will dramatically increase.

As mentioned, the action of moving a player who is under contract with a club to another club is called a transfer. This term is mainly used because the registration details are transferred from one association football club to another. The new club usually pays for the player's services. Thus, as compensation for the previous club losing the player, the new club pays them a previously determined amount of money – "transfer fee". This compensation can vary between straight monetary value or player exchange, with or without a monetary add-on. The player that is "offered" in exchange is in a difficult position since he does not have to approve the shift of clubs. Also, the club can reject to receive this player. Since 2002, UEFA established two fixed periods where it is possible to transfer players – called transfer window. The first one beginning before the start of the season, therefore from 1st July until 31st August, and the second taking place in the entire month of January. However, players without a contract with any team – called free agents – can be assigned at any moment of the season (Football-Stadiums.co.uk, 2020).

There can be different results in the procedure surrounding the signing of a player agreement. Such examples of results that have already occurred are:

1. Clubs that had a tentative agreement, but one of the parties dismisses from the final contract signing.
2. Clubs may consent to the move, but there is no agreement on the employment contract.
3. Finally, a dramatic situation in which two clubs complete the entire transfer process and the player dies unexpectedly.

According to Lukomski (2020), the events to procedure entirely with a player transfer are the following:

1. Agreement to move the player from one club to another
    a. An arrangement and a written deal to transfer the sporting rights of the player should be concluded between the former club and the current club
    b. This arrangement can be a financial amount or an exchange of players
    c. The move can be permanent or can go from half-season to two seasons via loan (there is no official maximum of seasons, however usually loans do not go over two seasons)
2. The approval of medical tests exercised by the new player's club
    a. This procedure is usually performed before the signature of the employment contract to make sure everything is fine with the player.
3. Signature of a contract of employment between the player and the new club. This contract includes:
        i. Wages
        ii. Duration of the contract
        iii. Signing-on bonus
        iv. Bonus payments
        v. Agent fee
        vi. Release/Buyout clause
4. Player's registration with the new football association
    a. This procedure is only needed when the transfer is made between clubs from different countries.
    b. The registration is a multistage procedure requiring the previous club to give the information and upload the transfer agreement and employment contract in the FIFA Transfer Matching System. After, the new club obtains the International Transfer Certificate (ITC) of the player. Finally, the player can be registered on the new football association.
    c. A working permit is also needed in some countries, such as in the English Premier League, where it gets difficult to obtain it in some cases.

The most singular and unique topic about a football player's contract is the possibility of including a buyout clause. This clause allows the player to cancel his contract (without having a just cause) with the present club by paying, typically, a tremendous amount of money. When the player completes this process, he becomes a free agent and can sign with another club. The buyout clause is not mandatory to include in the contract. It is only a possibility of agreement

when signing the contract. However, in Spain, this clause is indeed mandatory to include in the contract. Consequently, FIFA directives regarding this topic are not international law, and each country can decide independently (Madina Tatraeva, 2020). The commentary on FIFA regulations takes a more concise and direct explanation of how the buyout clause operates: *"The parties may, however, stipulate in the contract the amount that the player shall pay to the club as compensation in order to unilaterally terminate the contract (a so-called buyout clause). The advantage of this clause is that the parties mutually agree on the amount at the very beginning and fix this in the contract. By paying this amount to the club, the player is entitled to unilaterally terminate the employment contract. With this buyout clause, the parties agree to allow the player to cancel the contract at any moment and without a valid reason, i.e. also during the protected period, and as such, no sporting sanctions may be imposed on the player as a result of the premature termination."* (FIFA, 2006).

2.3 Football Finances

Nowadays, it is observed that money and interests behind wealth are more and more surrounding the football environment. The clubs are investing ridiculous quantities in players' transfer fees and agent's commissions. It is becoming more and more essential to have good managers to take care of the financial assets of the club, which are the players. Currently, a player can be worth more than a stadium or a training academy. Hence, it is almost mandatory to be maximum efficient in taking care of the club's assets.

According to Deloitte's annual study of the European football market size, the total revenue of European football continues to increase – the value of the season 17/18 was 28.4€ billion (Barnard et al., 2019). With a contribution of 14.7€ billion only in the top 5 leagues (England, Germany, Spain, Italy and France). Portugal was in the 9th position in this study, with overall revenue of 431€ million. It is possible to see this revenue growing substantially in the past years due to the broadcast deals. Clubs are receiving enormous quantities of these deals, and their leading receivables are from the broadcast. For instance, Real Madrid sold its sponsorship rights to a company for $224 million in a four-year deal. Also, Barcelona signed a shirt sponsorship for $246 million.

For those reasons and many others, Real Madrid and Barcelona are at the top of ranking the most valuable clubs in the world in 2019, made by Forbes. The top 5 of this ranking are: Real Madrid evaluated in $4.239 billion, followed by Barcelona with $4.021 billion. The rest of the top 5 are Manchester United ($3.02 billion), Bayern Munich ($3.024 billion) and Manchester City ($2.688 billion). On the rest of the top, essential to highlight that only clubs from the Big-five leagues are in the ranking, with most clubs from the English Premier League (Forbes, 2019a). Another example of the increased amount of money in the football deals is the tournaments' prize money. For instance, the prize money for the complete edition of the 2018-2019 champions

league achieved a record of $2.28 billion, more than 30% than the previous edition (Forbes, 2019b).

Another aspect to verify in the past years is that the overall player's transfer fees increase yearly. Mainly because of this growth in football investments. From the top ten highest transfer fees, the seven highest are from 2017 until 2019 summer. On the top are the Neymar's transfer (2017-2018) evaluated in 222€ million, followed by Philippe Coutinho's (2017-2018) with 145€ million. The last of these top ten is the transfer of Eden Hazard to Real Madrid for 100€ million in the 2019 summer (Statista, 2019). Figure 2 shows the increasing tendency of higher transfer fees across the last two decades. Furthermore, according to the 2020 ranking of the top100 highest-paid athletes globally, fourteen football players can be observed in figure 3 (Forbes, 2020). These high values are one more proof that players are one of the best assets of the clubs, and they must be well managed to provide the best club performance possible.

Figure 2 - Number of transfers per year achieving the all-time top 100 (adapted from bbc.com)

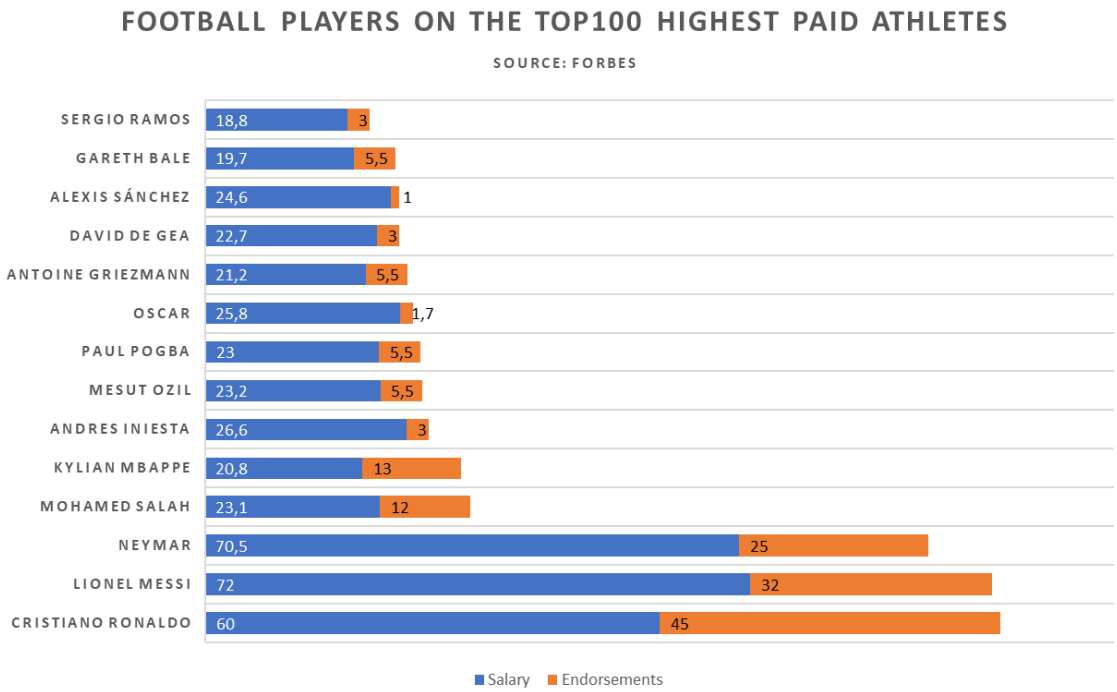

Figure 3 - Highest paid football players in 2020 (adapted from Forbes)

Analysing the TransferMarkt (accessed in February 2020) ranking of the most valuable squads, it is possible to see how high a team can be evaluated (note: the value of the teams in continuously on actualization). All the clubs in the top five had squads evaluated in more than 1€ billion. Manchester City was leading with 1.29€ billion and Liverpool FC next with 1.19€ billion. Finishing the top five were Real Madrid, FC Barcelona and Paris Saint-Germain with 1.08€ billion, 1.06€ billion and 1.02€ billion, respectively (TransferMarkt, 2020b). These values are constantly updated since the performance of each player affects their market value.

2.3 “It is all about performance” mentality

The sports industry is a results/performance-oriented industry, and football, of course, is included. Most of the spectators are not concerned about the financial efficiency of their teams. They only desire success in terms of sports results. As shown in the next chapter, Fizel & D’Itri (1997) showed that when the win ratio is considered to study performance, it is the factor that most influence a manager's success. Indeed, the central issue for fans in football is the performance of their team. However, if the team performs poorly, they will continue to support the team. Contrarily, to business companies where the typical consequences of a period of low performance are the loss of customers or a corporate reorganization, liquidation or acquisition by competitors, in football, all clubs retain a highly loyal fan base. These fans under no circumstances would support another team, independently of the results.

On the other hand, this loyalty of the fans increases the pressure of the only person on can be easily blamed – the coach. In other cases, also directors or players (Audas et al., 1999). Consequently, a problem arises with managing a football club. Sometimes it is not possible to be successful in terms of performance and be efficient in managing the actives of the club. So, a trade-off appears: the directors can either choose to invest in the best players to achieve the best performance and delight the fans or choose to contain costs and gradually rise on a progressive, practical project, which is difficult to accept as a fan. For instance, in a study taken by Haas et al. (2004), the authors concluded that the efficiency ranking and the league table was substantially different. One example was the difference between SC Freiburg and Borussia Dortmund. The first was better efficiently. However, Dortmund is a top table club, consequently in terms of performance behaved better.

2.4 Lack of studies regarding player turnover vs performance

The purpose of this dissertation is to measure the turnover in football teams in the six major leagues of the world and conclude about the impact of that turnover on their performance. According to the UEFA ranking of the countries in 2020, the top six leagues are followed by ranking order: Spain, England, Germany, Italy, France and Portugal (UEFA, 2020). Also, the topic of turnover in sports is not vastly explored. The ones that studied the issue were more focused on taking conclusions for public companies with football data than help to maximize the efficiency of football. Besides, studying the effects of performance in football is an easier job than analyse performance on regular companies. In a study taken by ter Weel (2011) – more explored in the next chapter – the author demonstrates some benefits of studying football to predict events in other industries. First, the performance outputs can be easily measured weekly. Next, those outputs are resumed in three possibilities – win, draw and loss. Thirdly, managers decisions and investments are public and observable, such as the number of players bought or values of each investment. Next, those decisions have an immediate and visible impact. Finally, football comprises one big homogeneous industry, which makes it easier to compare outputs between clubs. Also, the identification of problems is easier and less costly. The next chapter will go deeply on analysing the studies related with turnover either on sports or on general companies.

The authors that explored players turnover in sports went for approaches like task interdependence or the impact on the demand. Some examples are taken by L. Davis et al. (2014) and Morse et al. (2008). On the other hand, the measure of the turnover could be related to the variation of minutes played by the players. Consequently, the results could conclude about transfers of players and about variations inside a team and the impact of that changes on the performance. Another critical method is data envelopment analysis, and many studies are using this method to evaluate performance. It is an advantageous method of measuring efficiency using several inputs and outputs.

2.5 Football players demographs

The problem this dissertation is focusing on is indeed the stability of a football squad and its impact on the performance. In order to understand how relevant the problem can be, in figure 4, it is possible to look at the percentage of new players per club in the major football leagues – which are the focus of this dissertation.

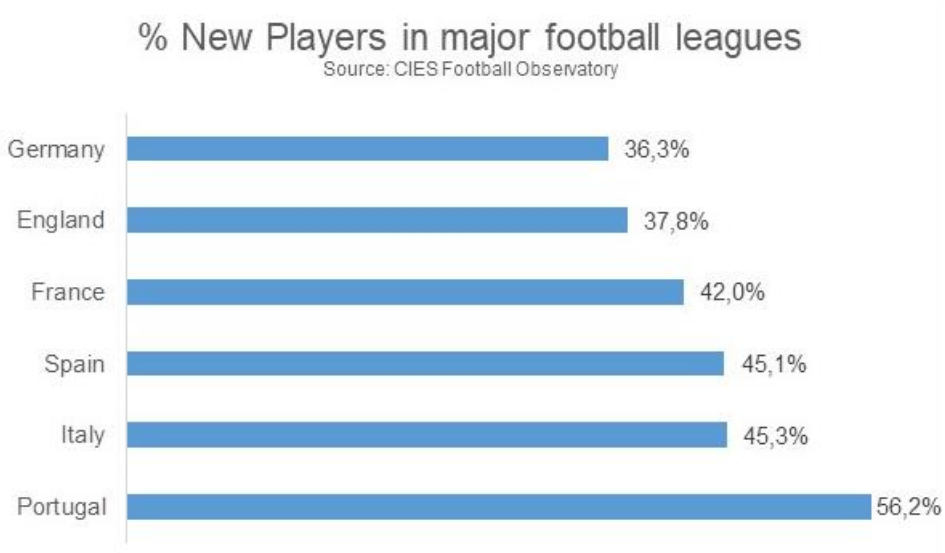

Figure 4 - % of players joined the first team after January 1st, 2016. (CIES, 2016b)

As observed, the league with more new players in the Portuguese league (NOS) with more than 50% of new players after 2016. Germany from the six leagues is the one with the lowest value of new players with 36%. (CIES, 2016b). Is this high percentage a problem for the performance of the team? That is the main question that will be answered in this dissertation. Besides the new signings, if the study included the players from youth academies that integrated the first team, the values would be higher. The average percentage of new players integrated into first teams rose from 41,2% in 2009 to a record value of 48,1% in 2016 – this means that approximately half of the squad is new from a year to another. To take a closer look at those numbers on each of our leagues in focus, figure 5 summarizes it.

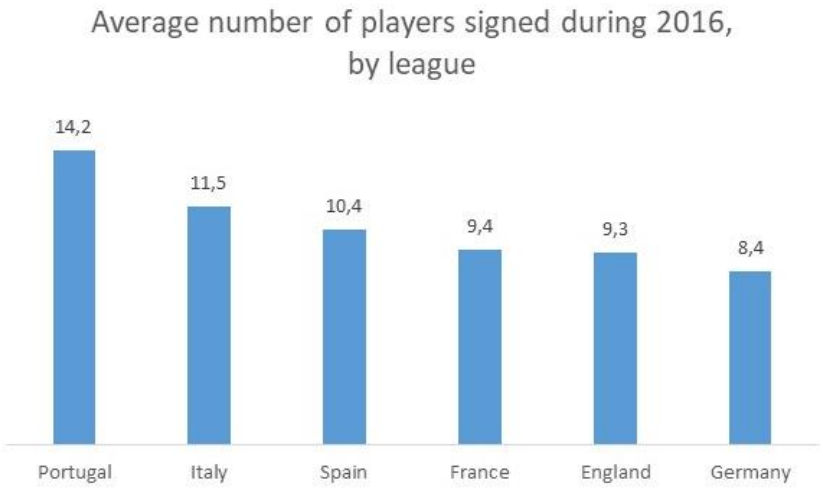

Figure 5 - Average signings in 2016 by league (CIES, 2016a)

Figure 5 shows that the average of the six significant leagues is around 10,5 signings in 2016, giving an idea of how turnover is a reality and how important it is to analyse it. In football, the clubs present themselves with 11 players on the field. Consequently, it is almost a new team in terms of signings (CIES, 2016a).

Nowadays, it is getting noticed that the average staying of players on the respective team has been declining in the past decade. Some examples are the average stay on Bundesliga and English Premier League. According to a 2017 report of ESPN UK, Bundesliga had an average of fewer than two years stay on the club per player – more specifically 1,95 years. Only six clubs had a stay over two years on the entire league. (Stephan Uersfeld, 2017) The case of the English Premier League is different. On average, the clubs of the English league retain their player for about 2,83 seasons (talkSPORT, 2015).

Consequently, with these two cases, it is possible to observe how important it is to study whether this turnover can impact the performance of the team or not. CIES Football Observatory, in an article of 2019, affirms that the average stay of UEFA leagues was decreasing from 2009 until 2017, where was achieved the minimum value ever of 2,22 years. From 2019, this value has not suffered many changes. In figure 6 below, it is possible to observe the year of recruitment and the respective percentage of players by 2009 and 2019 (CIES, 2019).

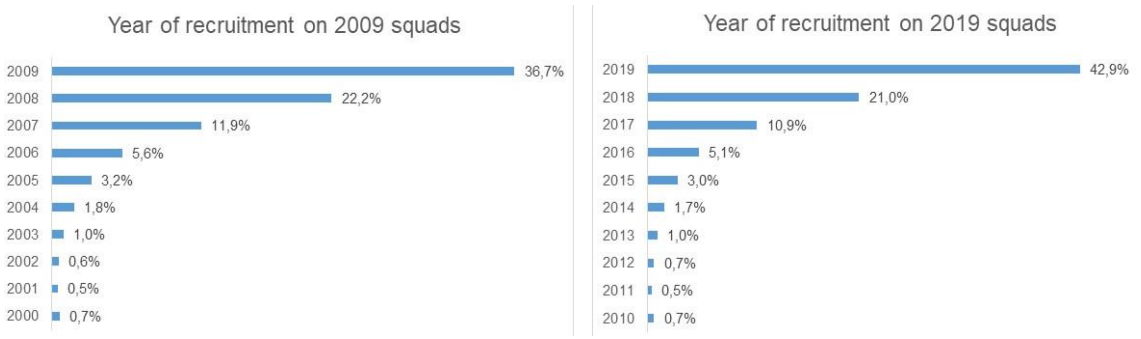

Figure 6 - Comparison between the year of recruitment of players in 2009 and 2019. Adapted from CIES monthly report 49

More than 50% of the squad was hired within the same year or previous year in both cases. In terms of signings in the current year of practice, 2019 achieved more than 40% of the entire squad.

3. Literature Review

The company's performance is highly dependent on how the workers and managers behave in terms of productivity and decision making. The same happens in football clubs, where the players or managers are essential for the team's performance. In sports teams, it is often difficult to know whom to blame for poor performance, mainly because the processes that result in the performance are not public, and the outcome is. Consequently, sometimes the managers are wrongly blamed. The main goal of this chapter is to go through a wide variety of authors to analyse whether there is any relation already studied between the squad turnover in the major football leagues. Next, it will be shown how the workers turnover could influence a company's performance (either sports company or business company).

The chapter will be divided into two main topics. First, some contextualization about job turnover is shown, with studies on job rotation, job turnover and manager turnover to have a general idea about the impact on companies' performance. Secondly, focusing only on sports, studies regarding manager and players turnover will be shown, which is the dissertation's focus. Finally, some studies regarding productivity, effectiveness, and efficiency are explored to clarify those terms better.

3.1 Job turnover in general companies

Intending to understand well the turnover in sports, let first analyse the impact, motivations and consequences of turnover on public companies. From the bottom workers to the CEO, all types of turnover influence performance one way or another. There are some critical studies regarding collective turnover which resulted in contributing with essential insights of the influence on companies' performance outputs such as customer service (Koys, 2001), financial performance (Batt, 2002; Huselid, 1995; Michele Kacmar et al., 2006) or labour productivity (Guthrie, 2001; Siebert & Zubanov, 2009). Another approach can also be held, which is studying what influences the collective turnover. Hancock et al. (2017) took that approach. They analysed the impact of recent prior firm performance on the turnover and the effect of contagious turnover, suggesting that contagion effects can happen at an aggregate level. The contagious concept related to turnover suggests that when workers observe their colleagues going on job search activities and accepting offers, they also start to feel available to do the same.

Hausknecht & Trevor (2011) also studied collective turnover through a framework that allowed to analyse the antecedents and the consequences of the collective turnover. Usually, this type of turnover is measured with separation, instability or retention rates. Also, the turnover rate is different depending on some factors: voluntary/forced turnover or the quality of the leaver (it can be a disruptive leaver with a significant impact on the company). Moreover, the consequences of collective turnover are likely to be negative. However, under certain circumstances, it can also

be beneficial. Heavey et al. (2013) similarly explored the antecedents and the posterior impact on performance.

On the other hand, Hancock et al. (2013) and Park & Shaw (2013) just focused on the performance impact. Hancock et al. (2017), to prevent the turnover, focuses on High Commitment HR Systems allied with trials on increasing satisfaction, commitment and perception for fairness levels to limit the collective turnover. Furthermore, another critical approach is investing in improving interpersonal team relationships.

Mohsin et al. (2015) explained the staff turnover in luxury hotels by some other antecedents. In this study, the feeling of job security, earnings and loyalty to the company were taken as variables. Similarly, the organisational enthusiasm and the feeling of a stimulating job were variables too. From these variables, the ones that showed no guarantees of increasing turnover were the high salary and job security. O'Fallon & Rutherford (2010) appointed as significant causes of staff turnover the treatment by superiors, excess of working hours, job pressure, scheduling, training, better opportunities in a different place, and physical demands of the job. Also, Hinkin & Tracey (2000) concluded that work environment and poor supervision outcomes in a higher will of leaving the job. With different results, Williams et al. (2008) showed that age, low unemployment rate and remuneration were crucial factors contributing to a higher job turnover.

Similarly, Ogbonna & Harris (2002) concluded that low remuneration is the crucial point for turnover. The most determinant reasons are related to retirement, illness, death, pregnancy or leaving the location of work. On the other hand, avoidable turnovers occur when the reasons are related to complaints with wages, lack of training, excessive work stress, relation with superiors, working hours and transportation difficulties (Mohsin et al., 2015). Another study taken by Mohr et al. (2012) also found a robust negative relationship between workers turnover and operational performance.

In football clubs, the player turnover rate can also be related to voluntary/forced turnover and the quality of the leaver. Imagining a player with several problems with directors and managers can be dismissed, or imagining a team captain is sold. Both situations can easily damage the performance of a team. Consequently, improving the interpersonal team relationships not only help in firms but also on football teams. Another integral approach and field of study is job rotation. There are many studies analysing job rotation inside big firms. The thoughts taken in these analyses can easily be compared to a change of role or position of a player inside the pitch. This change makes sense since it requires different responsibilities and decision making towards the same common goal of the team.

Nevertheless, what is job rotation? Kampkötter et al. (2018) define it as a lateral worker's transfer between jobs inside a company without changing the hierarchical or salary rank. Many companies choose to have this approach to keep the motivation of their workers on high standards or reallocate underperforming employees. The author concluded that the underperforming workers are the ones who are more often rotated between jobs. However, the

results showed that the rotation of high performers is less prevalent; it is beneficial for these workers for two years after the change.

To sum up, the ones that rotate more often are the ones that benefit the more minor, whereas the ones with more advantages do not rotate as often as they should (Kampkötter et al., 2018). On another approach related to job rotation, Sorensen et al. (2007) searched for explanations and determinants contributing to job mobility. For instance, the authors enumerated some economic conditions and societal characteristics – including war, monetary crises, technological breakthroughs, civil rights movements, labour battles, or even a tendency of corporate takeovers. Also, these determinants are included companies' staffing policies, career interests, the desirability of mobility or industry differences – like gender composition, wages, the intensity of work or perspectives of industry growth. Eriksson & Ortega (2006) showed the employee learning theory, which concludes that those who rotate more accumulate more human capital due to exposure to a more significant number of experiences.

Furthermore, from the firm's perspective, companies who rotate their workers learn more about them since the workers are observed performing different jobs. Also, rotation can be essential to stimulate the workers. Otherwise, they can fall into repetitive tasks and get bored and demotivated. Indeed, the re-allocation of human resources is another way of job rotation. However, the rotation can be costly for the company if the worker has no experience in the new position.

It is known that companies tend to make significant investments in human talent to create teams capable of taking them to a higher level. The importance of talent's management is becoming crucial to elevate companies' performances. Nevertheless, what is the impact of a change on the manager position in a company? Previously, some conclusions related to job rotation and job turnover on lower-level workers were showed. Next, it will be shown some significant conclusions.

Sports companies studies offer an easier way to evaluate and search for solutions in terms of organizational issues since its records are public and because sports companies are getting more similar to public companies. Most of the studies analysing the turnover of leaders in companies are compared with sports since, as explained before, it is easier to see the impact on companies' performance. Lieberson & Connor (1972) compared the influence in changing top managers (new presidents or board chairpersons) with the impact on the state of the economy, the industry where the company is inserted and the company's position inside the industry. They defended that a corporation's performance is reflected on the personal characteristics of its chiefs, like a sports team's success is credited to the coach's ability. However, when the performance suggests some troubles, the requests for justification fall on the leaders too. Gamson & Scotch (1963) defended that an organization in trouble is the one that leaders assume and show themselves to the blame. Lieberson & Connor (1972) concluded that corporate performance was less influenced by leadership and more by environmental or organizational issues. This study was continued years later with Weiner & Mahoney (1981), which explains the corporate performance

as a combination of environmental, organizational and leadership influences as the previous authors did. Each of these factors contributes differently depending on the dimension of performance that is being considered. So, the choices of the leaders are only a part of the whole responsibility of the company. Despite this, the leaders are always the easier way to blame someone if something goes wrong. It was seen that most of the studies suggest that manager turnover occurs when companies are misbehaving in terms of performance for a while. However, others have measured the firm structure (Denis et al., 1997), or take-over threats (Huson et al., 2001), or firm expectations (K. A. Farrell & Whidbee, 2003) or even the composition of top executive teams (Fee & Hadlock, 2004) to predict the firm performance.

In 1985, Reinganum studied the effect of executive succession on firms' stock prices to measure their performance. Depending on the company, the chief may have a different share of influence and power. Consequently, a change in this position can signify severe changes in the policies and strategies of a corporation. The mainline of thought when a managerial turnover happens is that change will improve performance. It is frequently showed in financial journals and might be called “common sense”. This line of thought is supported by the premise that individuals can control organizational issues by themselves. The way that performance can indeed improve is when the new chief executive avoids some of the mistakes done by the previous chief. However, this view, apart from some isolated cases, no systematic or empirical evidence supports it.

Another study topic is the one that compares the manager change to an outsider substitute or an insider substitute and its impact on the firm performance. Allen et al. (1979) concluded that an outsider substitute has a worse impact on the team. Also, Grusky (1963) found that insiders’ successions are less troublemaking than an outsider succession. The only two theories proved with empirical support predict a negative relationship between performance and succession, and others say there is no relation. Reinganum results only showed performance improvements on cases with an outsider new executive, a small firm, and with a previous chief that left the firm. However, not all cases of small firms with new outsider chiefs have experienced positive effects (Reinganum, 1985).

Finally, studies, which consider CEO turnover, find that companies’ boards are more likely to appoint outsiders substitutes when the CEO turnover follows a period of significantly low performance (Parrino, 1997). Farrell and Whidbee compared the results with the expectations and forecasts before the CEO turnover. Their results suggested that boards choose outsiders for the position when the firm’s 5-year forecast is no optimistic and does not suggest improvements. Also, when those forecasts are even more uncertain than the risk of appointing a new person to the CEO position. Companies’ boards are willing to take those risks to better forecast for improvements for the company future (K. A. Farrell & Whidbee, 2003). Farrell and Whidbee continuously said that one of the CEOs responsibilities is to control and influence the public’s perception of the firm, both to the press and public in general. Moreover, they also found evidence of the relation of the press pressure and CEO turnover. When the press increases the scrutiny of

poor firm performance, the likelihood of CEO turnover also increases. Finally, they admit that one primary responsibility of the CEO position is to control and manage the analyst's opinions and expectations (K. A. Farrell & Whidbee, 2002).

Another critical analysis of companies' performance was the investigation of Denis & Denis (1995). They compared different types of resignations, such as the impact of forced resignations compared to voluntary resignations on a company's performance. Thus, two scenarios were considered: firstly, managers voluntarily resign from low-performance firms; secondly, managers are replaced by corporate boards even if they are not responsible for their poor performance. The expectations were that neither of these scenarios would necessarily result in improving performance. A study taken by Bernard et al. (2018) put together the CEO "Outsiderness" with voluntary and forced departures to analyse the impact on performance. The conclusion was that turnover has a positive and significant impact after five years of change and that positive effect is higher when the new CEO is an outsider. Most forced resignations were not due to board monitoring but to other factors, such as block holder pressure, takeover attempts, financial distress or shareholder lawsuits. The results from Denis and Denis were expected to be as previous studies, where firms showed poor stock price performance on the periods after the manager turnover. Although, the study showed that firms tend to show decreases in performance in the last three years before the change and minor improvements after the change. They also could demonstrate that forced resignations and expected retirements had quite different results. First, the forced ones showed a large and significant decreasing performance before the change, and after that, it was considerable observable improvements. On the other hand, expected retirements showed that performance did not change much before the change and showed some minor improvements (Denis & Denis, 1995).

3.2 Manager turnover in sports

In the last two decades, people have been observing a transformation of sports teams into actual companies (or franchises). Like any other kind of company, they also focus on having the best results from their main assets, the players. Like big organizations, it is also necessary to have the best managers and coaches to make the team attractive. Any sports team that competes not only ends up having a similar goal but also a similar organizational structure. Therefore, sports teams also offer a starting point for study and performance comparisons through the skills of managers (Audas et al., 2002).

Moreover, manager turnover appears to be increasing substantially in the past years. Consequently, it becomes crucial to study the causes of this turnover. The study of the manager turnover started with Grusky's study in 1963. The author studied the managerial succession and organizational effectiveness with data of the MLB – Major League Baseball, the professional baseball American league. The study is based on an assumption of a common effect that the rate of succession is high. It would produce declining effectiveness on the organization level. On the

other hand, if the effectiveness is low, it would encourage a high rate of manager succession (Grusky, 1963).

According to Football World Rankings (2018), on the 16/17 season, all significant leagues had more than five sacked coaches - Serie A and Ligue 1 had seven dismissals, Premier League and Bundesliga had eight, La Liga had fifteen, and the league with most dismissals was Liga NOS with 16. Besides this, another statistical aspect of football coaches turnover is that 82.7% of the turnovers are forced – data from English soccer leagues - whereas on CEO turnovers, this percentage is 16.2% - data from large US firms (Audas et al., 1999; Huson et al., 2001).

The most popular and accurate conclusion regarding manager turnover admits that the changes in managers are a scapegoating ritual. Gamson & Scotch (1963) were the ones that advanced with this study, with an investigation on manager changes in baseball. The authors admit that this scapegoating ritual is more a way to appease the fans than increase performance. Besides Gamson and Scotch, also Brown (1982) and Grusky (1963) supported this theory. Through NCAA college basketball teams, Eitzen & Yetman (1972) reported that changes in coaches made none or little difference on team's performance. Brown (1982) concluded that the succession effect was barely noticed on NFL teams and admitted that a scapegoating view of succession is plausible. Grusky (1963) concluded that a change in leadership could provide a slide on performance due to tensions or instabilities. However, this slide on performance would only increase the probability of a leadership change again, and a cycle starts again. Allen et al. (1979) reported, through MLB teams, that the rate of managerial succession has a significant adverse effect on current team performance. Also, Eitzen & Yetman (1972) found that the organisations' coach turnover frequency and effectiveness were negatively correlated.

Ter Weel, in 2011, also tried to conclude whether the manager turnover improves firm performance with data from the Dutch Soccer. Indeed, and as mentioned before in this document, soccer is an easier way to make conclusions about firms' decisions. Not only because of both manager characteristics and decisions but also because firm outcomes are easily and quickly observed. Moreover, the fact that results information is so easily accessed provides managers to be sacked during the season. Also, the impact of turnover can only be seen after some years after the resignation, in which some other things, besides the CEO position, can also have changed (ter Weel, 2011). Despite some of the advantages that have already been said, it is plausible to enumerate some of the benefits of studying football to have conclusions about companies in general. First, the outputs and firm performance can be measure quickly and weekly, while business firms' performance is only measured through financial statements on a year-to-year basis. Secondly, the performance is easily and directly measured by a win, draw, and loss, making it a lot easier and simple. On the other hand, business firms rely on different measures to analyse performance: accounting, equity, and others. Consequently, the output performance can look differently under these several measures. Third, managers decisions and investments are public and observable since it is possible to see the number of players bought and the values of each transfer. Moreover, the impact of those decisions and investments become

effective immediately. In contrast, in business firms, the investments take more extended periods to become effective and positive. Fourth, the characteristics of the managers are available. Some of them are the ability and managing performance in the past, experience as a worker (in this case as a former player), the kind of player used to be (referring to offensive-defensive aspects) and the number of country representations as a player. All of these make a manager more complete and more capable. While, on firms, the information about CEOs characteristics is often unavailable and only known when the new person takes the position (ter Weel, 2011). Finally, football comprises one homogeneous industry, making it easier to compare results and show the performance between companies. Problem identification can become easier and less costly to observe and replace (Parrino, 1997).

ter Weel (2011) admits that the determinants for predicting and anticipating forced manager turnover on soccer companies are the number of players bought and their amount in investment. Moreover, the time remaining on the contract is also a determinant of whether the manager is thinkable to be sacked or not. When managers make a higher investment in players, they are more likely to be dismissed on a period of poor performance. The difference between one coach with highly talented players and one with weaker players is that highly talented players usually result in better results. However, there are many stories of coaches that transformed mediocre teams into an exceptional group of players, mainly changing few players. These stories suggest essential differences in the characteristics between coaches. The author ter Weel (2011) also concluded, if managers still have an extended period left on the contract, they are less likely to be fired because of the financial implications it would take for the company.

In clubs with a high tendency of managers turnover, coaches must show their knowledge, expertise and distinctive style quicker. On the other hand, when a manager leaves, what happens to the knowledge he acquires through time? Trequattrini et al. (2019) investigated the knowledge transfer between managers and the impact on performance. To achieve the desired results, the authors focused on five main variables: number of coach changes in the last five years, staff transferred with the coach, players transferred and investments in new players. Finally, the results showed that when a manager brings more than 50 per cent of his previous staff members, the team's performance is more likely to improve.

The results and effects of manager turnover on companies' performance, most of the time, depends on how high the corporate control on the decisions and act freedom of the manager is (ter Weel, 2011). In sports corporates, usually, this control is robust with the president present on match days. Consequently, this type of company often selects managers with style to adjust their will to change and implement new ideas. In general, firms that behave like that are the ones presenting better results in terms of performance. However, this is because few managers are willing to work in such a controlled environment where they cannot implement their ideas freely.

van Ours & van Tuijl (2016) put together data from 14 seasons about the number of coach dismissals on the top European leagues. The highest rate of sacked coaches happened in Serie A (Italy), with a mean of 8.4 per season. Secondly, La Liga (Spain) with 6.7 but closely followed

by Bundesliga (Germany) with 6.6 dismissals per season. Premier League (England) with 5.6 and Ligue 1 (France) with 4.7 complete the ranking. Team performance is the aspect that influences the most whether a coach is near to be fired or not. To measure team performance, the author prepared two indicators. Firstly, the cumulative number of points earned on the last four matches. Secondly, the difference between the expected results and the actual result is based on bookmaker odds. Generally, in football, the sensation given is that the performance is better after the dismissal of the head coach.

Consequently, and to have better accuracy, van Ours and van Tuijl created a control group where the coach was near to be fired, although the dismissal was not officialised yet – "coach dismissals that did not happen". Both dismissal and not dismissal cases presented improvements in performance after the sequence of bad results. Thus, comparing both results, the author concluded no causal effect of coach dismissal on team performance. uef

So why are coaches fired? The theory of a scapegoating routine of Gamson & Scotch (1963) remains plausible. Van Ours enumerated some possible reasons for the dismissal. First, a sacked coach can directly respond to pressure from unsatisfied stakeholders (e.g. shareholders, sponsors or many fans). Next, the pressure of the media, either online, radio, television or newspaper, are a severe factor to consider. Third, some conflicts between the coach and a board member, or problems with players, can be a cause. Next, the reason for most of the dismissals, the board blaming the coach for lousy performance, and the reality was that the bad results were only bad luck. Finally, the board can sack the coach to hand on the problem justifying that "doing something is better than doing nothing" (van Ours & van Tuijl, 2016).

In the studies of analysis of the determinants of the dismissal of coaches, it is important to emphasize some authors. Basically, it is almost unanimous that recent past performance is one of the essential determinants related to coach dismissals. Besides this, Audas et al. (1999) also found that the age of the coach matter for the decision of dismissal. Likewise, Dobson & Goddard (2001) concluded that most dismissals are due to variation in the league ranking and coach's age. Moreover, Salomo & Teichmann (2000) found that intensive media coverage of the period of bad results is determinant. Tena & Forrest (2007) discovered that the pressure of being close to the relegation zone is more determinant than the recent bad results.

Furthermore, another determinant is the managerial efficiency measured by the difference of places on the club's current position on the league ranking and the club's position on the budget ranking. Consequently, if a club is spending much, its position in the league is not higher than in the budget ranking, which increases the probability of a coach dismissal. At the same time, Bachan et al. (2008) concluded that the most critical factor is the place in the league, and neither of the coach's characteristics is essential. Another study taken by Barros et al. (2009) relates the coach's salary with the probability of being sacked and concluded that manager of expensive teams tends to be sacked earlier. More salary does not mean a higher probability of surviving in the position. Frick et al. (2010) showed that the salary of both the manager and the team wage

bill is critical in increasing the probability of a dismissal occur. The most popular ways to measure team performance are the number of points earned or the goal difference.

Almost every study proved that a coach dismissal does not improve the club performance. Poulsen (2000) and Salomo & Teichmann (2000) are other two examples. Dobson & Goddard (2001) concluded that the performance after the change worsened, so the dismissal is considered disruptive quickly. Koning (2003) made an exciting approach to comparing the new manager's performance with his predecessor. However, the author compared the quality presented on matches against the same opponents, with different managers, resulting in a much worse performance of the new manager. Muehlheusser et al. (2016) explored whether the heterogeneity of a team together with managerial change influences team performances. The results showed that the performance only enhances in homogeneous teams – the abilities of the weakest players on the field are similar to the strongest ones on the bench.

Shook (2016) explored the impact of a coach dismissal either on short and long-term performance and compared the actual performance to the previous expectations. One example is the study of d'Addona & Kind (2014), where the conclusion was that past performance is the most determinant factor to the dismissal, as mention in most of the studies. However, the authors concluded that the period of most influence on the decision is the previous two weeks. Shook (2016) took the evaluation of the team performance to another level. Instead of a win/draw/loss register, the author measured performance with pre-game betting odds and compared it afterwards. This method allows having a better assessment of the opposition calibre. Finally, the author concluded that performance matters and the age and nationality (foreign or native coach) could influence dismissal. There is some evidence that foreign coaches are more likely of being fired. Furthermore, the players' contracts are becoming more and more expensive, consequently sacking the manager is the more economical response to lousy performance in contrast with buying better players.

Audas et al. (1999) were pioneers in studying team performance and the dismissal being voluntary or involuntary on football clubs. The dismissals are considered voluntary when there is clear evidence that the manager took the decision. That evidence includes an instant transition to another team with at least the same level, assume the position of national team manager, or simply a voluntary retirement of football. On the other hand, for involuntary dismissals, the study concluded that the most determinant factor is the win ratio of the nine previous matches. Moreover, whether the position in the league is higher than when the coach took charge is also significant. Age is essential too. However, no other managerial human capital characteristic is. The latest performance is a factor for voluntary dismissals, too, however, with much less impact. Despite most manager job offers are for recently successful coaches, the willingness to accept an offer is increased by a streak of poor performance (Audas et al., 1999).

Scully (1994) used a method of measuring managerial efficiency and observed that job survival is directly proportional to managerial efficiency, which was predictable. The efficiency was measure by the residuals from a stochastic frontier regression of the team's win ratio and its

ratio of points. This efficiency is related to the capacity of the manager to allocate the team's recourses to the field to maximize the probability of winning. Consequently, a run of lousy performance results can be explained by the lack of talented players and poor management. Another study of managerial efficiency related to turnover, either voluntary or involuntary, was taken by Fizel & D'Itri (1997). Through a Data Envelopment Analysis (DEA), the efficiency is measured with relations between win ratio and playing talent. In cases where the win ratio is taken out, it is possible to observe that both playing talent and manager efficiency are related to involuntary termination. However, when the win ratio is added to the study, the organizational productivity is resumed only by the question 'did we win?'. Basically, the author concluded that the manager is only blamed for winning or not winning. Consequently, a manager can be fired for not winning despite being efficient.

But how about the players' performance? Earlier in this dissertation, the impact of a change in the manager position who chooses the group that is more able to go for the match and in firms who make decisions directly impact the company's performance. Now it is time to analyse the impact of a change in the squad. As mentioned earlier, the players are the most critical asset of a club. There is a lack of analysis on this impact; however, with studies regarding the importance of measuring team performance, it can be possible to analyse the importance of players on the club's overall performance. The analyses of the impact of a change on the core level of the squad on the team performance will be next shown.

3.3 Squad stability, player turnover and team performance

As mentioned before, few articles, dissertations and studies regarding the relationship between stability and success on football teams. One of the few ones was performed by the CIES football observatory in 2018 – however, the study took data from 2009 until 2017. The conclusion was that the best performing teams were the ones that had a more stable squad than the least competitive clubs. The best performing teams had a percentage of new signings of 31.3%. On the other hand, the least competitive had around 42%. There were some league champions with a high percentage of new signings, although they were in leagues of countries where squads are traditionally unstable. The record percentage among champions occurred in 2011 with PFC Ludogorets Razgrad of Bulgaria – 92% of new signings on the squad. On the top ten in the table of champions with the lowest percentage of new signings are many clubs of the five major leagues, except for Italy (Serie A):

- FC Bayern Munich (Germany – 2016): 9.1%
- FC Barcelona (Spain – 2012): 11.5%
- Manchester United (England – 2010): 12.9%
- LOSC Lille (France – 2010): 13.6%
- Chelsea FC (England – 2009): 14.8%

These numbers are especially remarkable when compared with the average percentage among champions of the major leagues from 2009 until 2017:

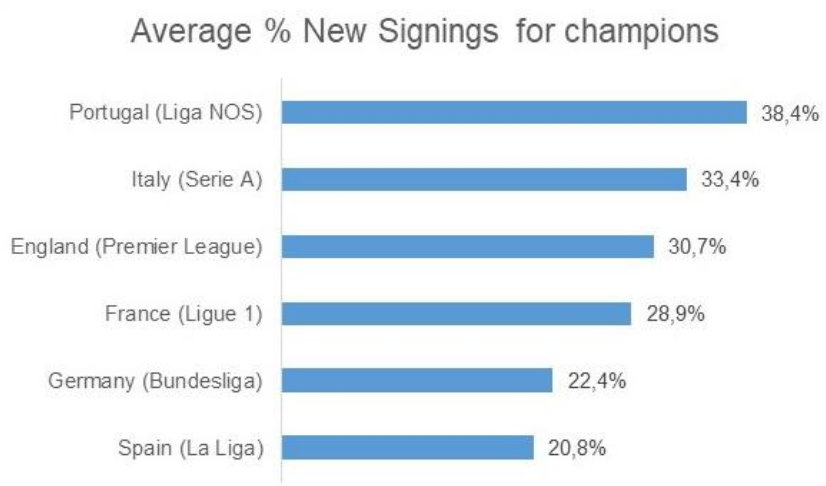

Figure 7 - Average % New Signings for champions on significant football leagues (2009-2017)

The average on all champions of all leagues was 34%. According to the authors, this number allowed to confirm that clubs with a stable squad are a majority (CIES, 2018). From this study taken by CIES, it can be concluded that performance and player turnover are negatively correlated. The most successful teams usually have lower turnover rates. Brown (1982), similarly, has proven that performance is negatively affected by the rate of player turnover and outsider successions affect performance by increasing the rate of player turnover.

One extremely relevant approach to measuring player turnover was the one taken by Morse et al. (2008). The authors studied the impact of turnover on-demand in the NBA – National Basketball Association. The primary variable of interest – roster turnover – was evaluated with a percentage of players on the team that played at least 60% of the games and were not on the roster on the following season. Statistically speaking, in NBA, teams lost 27 per cent of their players each season. In addition to this variable, the authors included others that can be interesting to evaluate performance also:

- Salary turnover: evaluates the salary bill that is not on the following season. It is a way to weigh players based on their salary, which is usually proportional to their quality.
- Previous year's attendance: despite this variable is not related either to performance or turnover. It can be used to analyse if the team is getting well supported or not, modifying the players' motivation.
- Current season winning percentage
- Previous winning percentage: can be used as a comparison

- History: meaning the number of championships the team has. This number is related to the pressure to win.
- Population: city population can be a crucial factor of pressure to the team.
- New arena: whether the team has a new arena or not, it can influence the supporters to come

In the same way as the previous authors, Kounetas (2014) studied efficiency related to Greek clubs' performance before and after the Euro 2004. The study is divided into two analyses: first, a bootstrapped DEA to determine the efficiency and, secondly, an investigation to determine possible factors affecting the efficiency. The bootstrapped DEA included input variables: the total of players' transfer expenses and contract renewals, general operational costs, clubs' total expenses. On the output variables, there are included: points earned and total attendance. Furthermore, other variables were considered: profit margin, financial exposure intensity (total assets to debt ratio), club's age and location. Also, Haas (2003) evaluated the efficiency of English football teams with the DEA approach. The author used as inputs the club's wages and salaries except for the head coach salary, which is used as a separate variable since the coach significantly influences team performance. Also, the size of the squad was included since it is directly related to total wages and salaries. On the output area were the point awarded by the team together with whether a team is included in international competitions or not and team's commercial output directly related with revenues. On the other hand, from the supporters' point of view the champion is always the best, in fact even the successful teams can be caught in less efficient periods. The results showed that using playing talent and coach capacities as inputs and points awarded and revenues as outputs, together with the population of the club's hometown, one-third of the team are being sufficiently efficient. Within the same line of thought, Haas et al. (2004) studied the efficiency of the German football teams by the DEA approach. The authors' results showed that efficiency ranking is not correlated with rank in the league, with SC Freiburg beating Borussia Dortmund on efficiency.

L. Davis et al. (2014) have taken an innovative approach, including the influence of task interdependence and turnover, to evaluate the performance of NFL teams – the American Football League. The authors concluded that in teams with high interdependence, turnover is even more negative on performance. In fact, most turnover cases are related to those teams, and when turnover happens, the team's processes are disrupted, norms are changed, new ideas are introduced, resulting in lower performance. Cohen & Bailey (1997) defined the concept of a team as a group of individuals working as independent entities inside an organization fighting for a common pursued outcome. However, organizational researchers use the term team for groups that have included high interdependence, which is a term related to the coordination of team members to achieve the desired outcome. Interdependence is as high as the necessity of coordination between team members to accomplish a task (Gully et al., 1995). When this coordination is optimized, the performance is efficiently maximized. Team stability also becomes a factor with such importance when talking about interdependence. Moreover, it can even be

more important than the overall individual talent of the members (Gammage et al., 2001). L. Davis et al. (2014) suggested that future research include variables such as quality of replacement, the experience of the replacement, experience, and the current team members or training procedures upon entry.

Carmichael et al. (2014) studied how to measure production and efficiency on Italian Serie A. The authors focused on two approaches: econometric stochastic frontier and data envelopment analysis. The first is based on regression analysis, and DEA is, as explained before, based on axiomatic proprieties and mathematical programming techniques. Boscá et al. (2009), also with a study on Serie A and La Liga, concluded that defensive efficiency is more critical in Serie A than offensive efficiency. In contrast to Boscá et al. (2009), the results showed that Serie A is more necessary to have better offensive performance than defensive, which is odd because Serie A is historically seen as having a defensive mindset.

In contrast, in La Liga, the opposite is true. Team success is easily measured by victories, reflected on points won and positive goal difference. Furthermore, the number of goals is related to the number of effective attacking moves – passing accuracy, possession percentage, shots on goal, and more. In the same way, defensive efficiency is related to the number of goals conceded and defensive skills and opponent attacking skills. The model is used as inputs to produce league-level success, attacking, constructive and defensive playing performance.

3.4 Productivity, Effectiveness and Efficiency

Productivity, Effectiveness and Efficiency are terms directly related to evaluating performance either on companies or sports teams. The correlation between performance and effectiveness started with Grusky (1963) with a study on the managerial succession on MLB – Major League Baseball. The author describes effectiveness as a concept related to an organisation's capacity to complete and succeed in its official aims. Other authors also explore this concept as Mouzas (2006). The author defines efficiency as a possible barrier or condition depending on the company's operating margins. It also defines effectiveness as the company's success in making one sustainable growth in earnings in all exchange relationships in the marketplace (Mouzas, 2006).

A manager can be efficient without being effective. Consequently, it can be confirmed that efficiency is not a measure of success but a measure of the operational distinction and productivity. In general, this term is about minimizing costs and seek for better operational margins. On the other hand, effectiveness is about its strategy to overcome and generate sustainable growth in its business environment. Companies invest in the best actives possible to obtain sales revenues and compensate the investment done. Companies seek to find business opportunities to generate a return on their assets with the aim of offsetting their financing cost. However, those assets are not free of charge since they always have the opportunity cost of

capital. This cost is related to choosing one asset instead of another or the cost of investing in other businesses opportunities of similar risk.

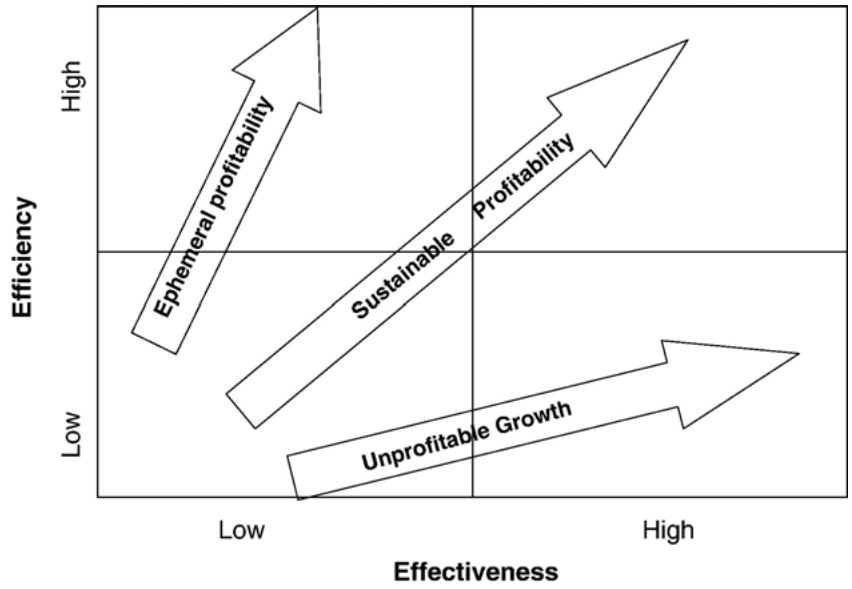

Figure 8 - The effect of different levels of efficiency and effectiveness adapted from Mouzas (2006)

In figure 8, it can be observed that a company should balance between efficiency and effectiveness to have the best profitability possible. If the tendency is to have more efficiency than effectiveness, the company will likely fall in a quick and non-durable profit. On the other hand, if the company have the opposite may fall into ‘unprofitable growth’ in the case of the opportunity cost of capital being higher than the resulting profit. The perfect balance of both terms requires companies to make more vigorous efforts in business networks. The author concludes that the first case means the company is neglecting the innovation and development of new forms of value to maintain profitability. Many companies usually focus on maximizing profit. However, they forgot to access the full impact of not investing in the growth through marketing and R&D (Mouzas, 2006). These decisions about efficiency and effectiveness are taken in some part by managers. The manager is one crucial part of the leadership of a company or team. They control and take important decisions to have the better response possible to problems shown.

Regarding production, which is defined as a process where the inputs are identified, prepared and transformed into outputs. In addition, the production unit that allows that transformation is denominated as an independent decision unit (DU). The production process can be observed in figure 9.

Figure 9 - Production process and DU

Assuming that all industries have limited input capacity, the quality of the production process, i.e., the efficiency of resource utilization and the maximization of products, must be studied. According to Cullinane et al. (2004), this issue has turned the concept of performance into a very significant definition. The principles of productivity and efficiency are fundamental to performance evaluation methods**.** Cullinane et al. (2004) defined DU productivity as the ratio between output and input and efficiency as a relative term that can only be determined by comparison or benchmarking. Before Cullinane et al. (2004), Vincent (1959) also defined productivity as the ratio of output and input.

Moreover, Mbuvi et al. (2012) defined effectiveness as the proper measures to apply, while efficiency is the correct way to apply those measures. Finally, efficiency can be divided into technical efficiency, scale efficiency and allocative efficiency. In addition, when multiplied by each other, the technical and allocative efficiency form the economic efficiency. This concept represents the optimum allocation of all factors in production resulting from the minimization of inputs to the level of output or the maximization of output at the level of inputs (M. J. Farrell, 1957).

First, technical efficiency is characterized as relative productivity over time, space or both. In an economic sense, the production frontier – linked to the maximization of outputs at a certain level of inputs – and the cost frontier – related to the minimization of inputs at the desired output level – determines technological efficiency as output-driven or input-driven. Second, scale efficiency relates to the divergency between the present state of production and the ideal production level. Finally, the allocative efficiency focuses on the production cost considering the information price and can be divided into profit maximization or cost minimization.

In figure 10, it is possible to analyse the inputs and outputs approaches, both only with 1 unit – variable scale incomes (a) and regular scale incomes (b). The curve *f* is the production frontier on a specific period for a given activity, and this curve can also be called the efficiency frontier. Technical efficiency is based on the distance from the DU to the frontier – B and D are technically efficient; however, P is technically inefficient. In both cases (a and b), the productivity of P is measured by reference, so the CP/C0 ratio gives the productivity of P. The reason P is inefficient is that it is possible to get a greater output with the same input, i.e. to hit point D. On (a), the technical efficiency for the output approach is given by the ratio CP/CD, while for the input is AB/AP. On the other hand, both for input and output, on (b), technical efficiency is given by the ratio AB/AP = CP/CD for any weak point.

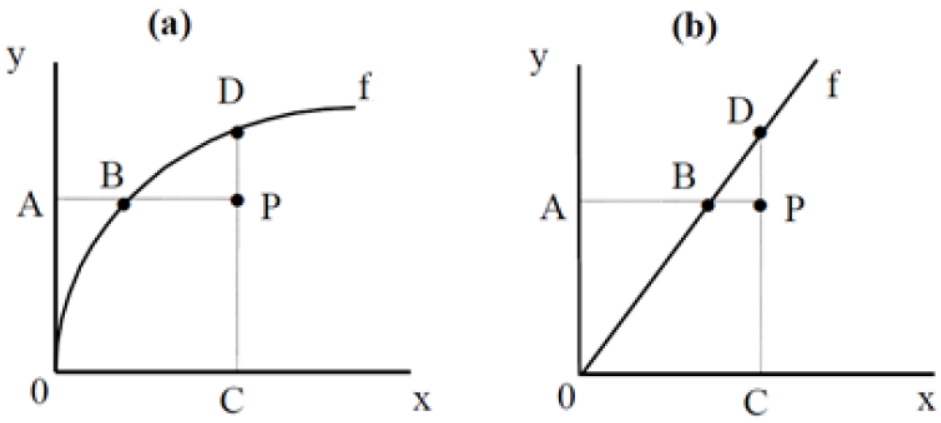

Figure 10 - Variable and Constant Incomes example (Coelli, 1996)

Economic efficiency is a result of technical and allocative efficiency. For this example, it was considered two inputs (x_1, x_2) to the same output level (y) – as observed in figure 11.

In figure 11, curve f represents the efficient boundary, and c represents the price ratio between the inputs. Regarding point B, B' is technically efficient, although it is inefficient concerning allocative efficiency. At the same time, point B is not only inefficient for allocative but also for technical efficiency. Allocative efficiency can be determined by the OB"/OB' ratio, while OB'/OB determines the technical efficiency.

The latest examples explained that efficiency and productivity were made based on a static perspective. However, if the examples were based on a dynamic point of view, it would be possible to observe a change in the position of the efficiency frontier. This change is mainly due to the industry's technological evolution, and as a result, the company may not be able to follow the dynamic movements of the market.

Figure 11 shows two frontiers in different moments of time, frontier f and f' on periods T and T-1, respectively.

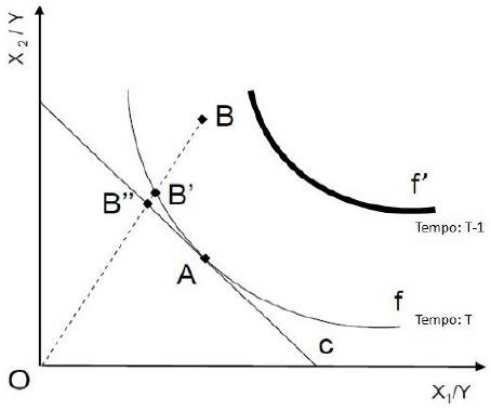

Figure 11 - Technical and Allocative Efficiency (Carvalho & Marques, 2007)

Frontier f DUs have achieved the same outputs with fewer inputs, i.e. those DUs have adapted dynamically to market progress. Those who remained in the same position at the border f could not invest in growth in the direction of the market movement.

In figure 12 is analysed the level of inefficiency caused by pure technical inefficiency and by scale operational inefficiency. From an input point of view, point B reveals two types of inefficiency: pure and scale. Curve f shows a pure, effective frontier and point A displays an optimum scale function.

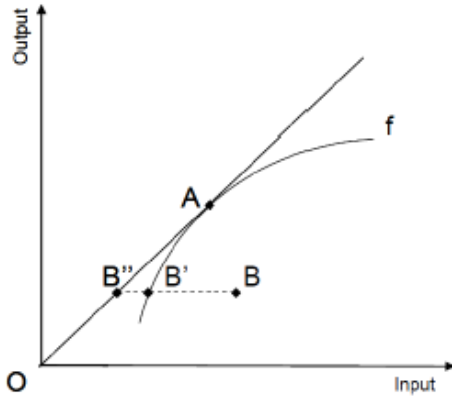

Figure 12 - Pure Technical efficiency and scale efficiency (Carvalho & Marques, 2007)

3.5 Chapter conclusions

To conclude, this chapter saw various important topics to contextualize this dissertation position in the existent literature. Firstly, it is shown an overall contextualization regarding three definitions that are explored across the dissertation. Subsequently, introducing the topic of turnover it is first shown how the topic was explored regarding public companies. Finally, some studies more connected with the sports world were explored. By exploring those studies, it was possible to conclude regarding the vast range of justifications and reasons for turnover in general. Additionally, by analysing how the primary authors explored the topic, it was possible to collect crucial insights for deciding which variables and datasets to use, and how both would be managed and studied. The last topic explored was especially crucial to conclude which variables to use in this masters' dissertation.

4. Methodology

This chapter goes through a theoretical background of the methods that will be used in this dissertation. In order to have more consistency and more robust results, this dissertation will use two methods of different types: Data Envelopment Analysis and Linear Regression.

4.1 Performance Evaluation Methods

These evaluation methods can be separated into two main groups, depending on whether it is a parametric approach or not: parametric approach and non-parametric approach (Drake & Simper, 2005). Both can be further subdivided into methods that use the construction of an efficient frontier, following best practice, and methods that do not use the definition of an efficient frontier, taking into account average adjustments (R. Marques & Silva, 2006). A more schematic way to organize this hierarchy can be observed in figure 13.

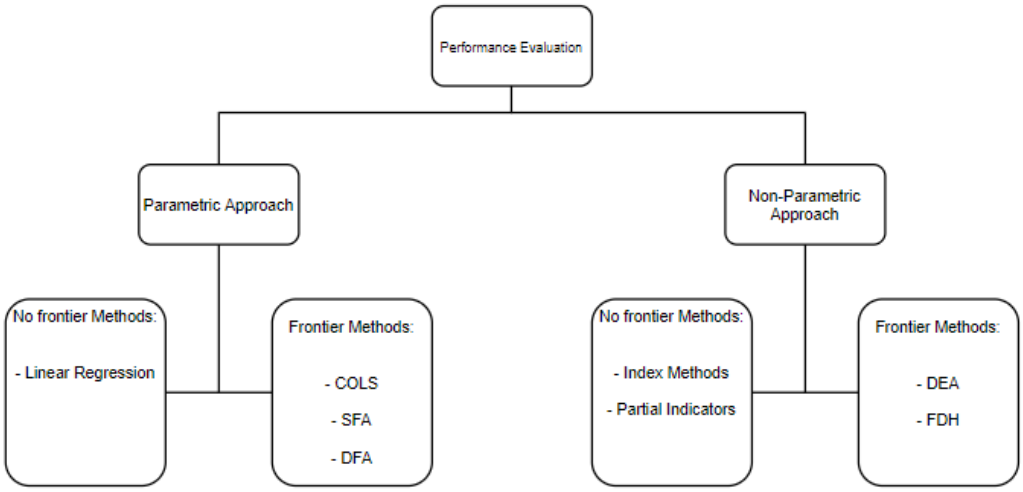

Figure 13 - Performance Evaluation Methods adapted from (R. C. Marques, 2011)

According to Mortimer & Peacock (2002), the methods that use the definition of an efficient frontier are the ones whose efficiency results obtained reflect the difference between the observed performance and the potential one. Furthermore, these methods evaluate performance on different dimensions since it is possible to study multiple inputs and outputs simultaneously.

The main goal of a performance evaluation methodology is to measure the efficiency of the different DU (decision units) of the study compared with the boundary. In order to apply one of the previous methods, first, an objective must be defined, typically from one of two types: output maximization from a given level of inputs or input minimization to a level of outputs. Secondly, the

efficient frontier must be estimated, to which the efficiency measures of the DUs will be compared. This step is where the first differences between methods start to reveal (Drake & Simper, 2005).

Methods with parametric approaches, such as linear regression, COLS (corrected ordinary least squares), SFA (stochastic frontier analysis) and DFA (deterministic frontier analysis), require the estimation of the function to allow the definition of the production frontier and the measurement of errors. On the other hand, non-parametric approach methods, where DEA is included, are considered empirical methods, which means that many data observations are needed to make it possible to estimate the production frontier of the DUs. In both cases, the efficiency ranking is determined by comparing the best practices or a medium adjustment (R. Marques & Silva, 2006).

For methods with a parametric approach, those using the production boundary can also be divided into two categories: stochastic methods and deterministic methods. The first, which considers random production, view the distance between the DU and the border due to random disturbances of the process. These methods may or may not include a parcel for inefficiency isolation. On the other hand, the deterministic methods consider the distance between the DU position and the frontier due to the DU's inefficiency (Simar, 1992).

Usually, approaches that use the output frontier are more reliable than those that do not use it. However, as far as parametric and non-parametric are concerned, there is no consensus opinion on the literature (R. Marques & Silva, 2006).

4.1.1 Data Envelopment Analysis – DEA

As stated earlier in this dissertation, DEA is a non-parametric technique for evaluating the efficiency of DUs with multiple inputs and outputs (Cullinane et al., 2004). On the other hand, Cook & Zhu (2005) indicate that DEA is an approach to assessing the efficiency of DUs, more precisely, on how inputs are transformed into outputs. As a result, the efficiency of the DUs is nothing more than the ratio of the weighted sum of inputs to the weighted sum of outputs.

The definition of Data Envelopment Analysis (DEA) was presented in the journal literature by a widely influential by Charnes et al. (1978). This paper established Edward Rhodes with a PhD, under the supervision of W. W. Cooper. The development of research leading up to this paper appeared to be overlooked in later literature. However, researching this diffusion of ideas can provide helpful insights into research problems that are still unexplored and insight into the research process. Farrell's seminal paper posted in 1957 regarding performance principles and their computation is a natural starting point. The richness of the concepts discussed in Farrell is evidenced by the fact that innovations over the next two decades have been focused on aspects and ideas (M. J. Farrell, 1957). The sources of the critical inventions are established, and the similarities to Charnes, Cooper and Rhodes are explored. Charnes et al. (1978) described DEA as a mathematical programming model applied to observational data, which offers a new way of obtaining empirical estimates of relationships – such as output functions and/or efficient production potential surfaces that are the cornerstone of modern economics.

DEA is a technique that evaluates the efficiency as a relative efficiency of the total DUs on the sample in the study (Cullinane et al., 2004). The weights of each variable are obtained by fractional programming to optimise each DU's efficiency. All DUs considered effective have a value of 1 and together form an efficient frontier, and the distance from the DU to the frontier is proportional to the increase in inefficiency (Chen et al., 2010; Hu et al., 2010).

As stated earlier on the subject of performance assessment methods, DEA may have two distinct orientations: it may have an orientation towards minimizing inputs to achieve a specific output level or optimizing outputs from a particular input level (C. Barros & Athanassiou, 2004). Both of these techniques can be used in two separate DEA methods: CRS and VRS. The first approach can also be called the CCR model in honour of its authors – Charnes, Cooper and Rhodes – (A. Charnes et al., 1978). Like the previous one and for the same reason, the second one can be named the BCC model – Banker, Charnes and Cooper (Banker et al., 1984).

According to Thanassoulis (2001), the DEA application requires some preconditions before beginning the analysis:

1. The entities of the study must be homogenous and with identical objectives and processes
2. The DUs in the study should operate in similar operational environments
3. The inputs and outputs must be the same across all DUs in the analysis

Wu et al. (2010) considered one of DEA's significant advantages because it also makes it possible to infer the number of inputs that should be reduced and/or the number of outputs that should be increased in place to ensure the DU inefficient for efficient units. Nevertheless, DEA has also some limitations, according to Sarafidis (2002):

1. Increasing the number of variables in the analysis only makes higher the possibility of having efficient DUs
2. Evaluates the relative efficiency instead of the absolute efficiency
3. Considers as inefficient all the DUs out of the frontier
4. Do not allow test hypothesis formulations

4.1.1.1 DEA – CRS Model

As mentioned earlier, Charnes, Cooper and Rhodes developed this approach in 1978 and considered the CRS technology and the overuse of variables. In order to optimize every last DU, fractional programming is used to assign weights to the inputs and outputs of each DU in the analysis. Consequently, each DU can have different input and output weights related to its maximization. Even if applied to other DUs, these weights can never be valued above 1 in terms

of efficiency. Therefore, the relative efficiency index focuses on comparing inputs and outputs of all the DUs included.

In figure 14, it is possible to observe a group of DUs classified by the DEA-CCR model, where only point C is considered efficient.

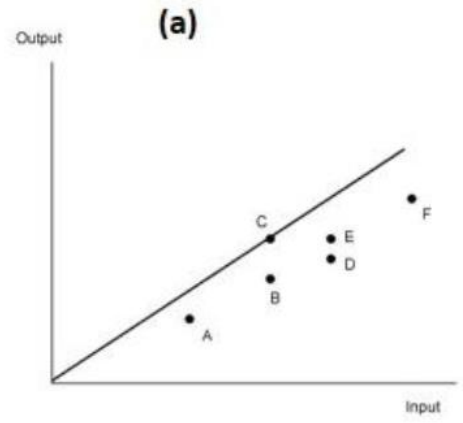

Figure 14 - DEA-CCR model example of production frontier (Kim & Harris, 2008)

In figure 15, there is the formulation for a CRS model on both input and output approaches. This general formulation assumes a *n* number of DUs to be evaluated, which consume x_{ij} inputs and y_{rj} outputs. Observing figure 15, the CRS model is formulated on both input and output approaches. An *n* number of DUs, which consume x_{ij} inputs and y_{rj} outputs, are assumed to be evaluated on this general formulation. The efficiency of DU_0 is defined concerning all DU_j (with j = 1, …, n) efficiencies in the study. Regarding the rest of the indexes, u_r and v_i correspond to the weights of either the outputs or inputs, and y_{ro} and x_{io} are the outputs and the inputs of DU_0 (Cooper et al., 2006).

DEA – CCR input-oriented	**DEA – CCR output-oriented**
$\text{Max } h_0 = \frac{\sum_r u_r y_{ro}}{\sum_i v_i x_{io}}$	$\text{Min } h_0 = \frac{\sum_i v_i x_{io}}{\sum_r u_r y_{ro}}$
With the constraints:	With the constraints:
$\frac{\sum_r u_r y_{ro}}{\sum_i v_i x_{io}} \leq 1, j = 1, \dots, n$	$\frac{\sum_i v_i x_{io}}{\sum_r u_r y_{ro}} \geq 1, j = 1, \dots, n$
$u_r \,\&\, v_i \geq 0 \;\forall\, r, i$	$u_r \,\&\, v_i \geq 0 \;\forall\, j, i$

Figure 15 - DEA formulation for CRS methodology (Cooper et al., 2006)

Nevertheless, after a while, it was identified a significant handicap on this formulation: the model generates infinite optimal solutions (Abraham Charnes et al., 1984). This downside has led to the emergence of a new model, the primal model. The difference is that instead of using fractional programming, it uses linear programming, as it is possible to see below in figure 16 (Cooper et al., 2006).

Primal Model

DEA – CCR input-oriented	**DEA – CCR input-oriented**
$\text{Max } z = \sum_{r=1}^{s} u_r \, y_{ro}$	$\text{Min } q = \sum_{i=1}^{m} v_i \, x_{io}$
With the constraints:	With the constraints:
$\sum_{i=1}^{m} v_i \, x_{io} = 1$	$\sum_{i=1}^{m} u_r \, y_{ro} = 1$
$\sum_{r=1}^{s} u_r \, y_{ro} - \sum_{i=1}^{m} v_i \, x_{io} \leq 0$	$\sum_{i=1}^{m} v_i \, x_{io} - \sum_{r=1}^{s} u_r \, y_{rj} \geq 0$
$u_r \,\&\, v_i \geq 0$	$u_r \,\&\, v_i \geq 0$

Figure 16 - DEA Primal Model formulation for CRS methodology (Cooper et al., 2006)

This first formulation in figure 16 makes it possible to formulate the corresponding dual model, visible in figure 17.

Dual Model

DEA – CCR input-oriented	**DEA – CCR input-oriented**
$\theta^* = \text{Min } \theta$	$\theta^* = \text{Max } \theta$
With the constraints:	With the constraints:
$\sum_{i=1}^{n} x_{ij} \, \lambda_j \leq \theta x_{io} \, , i = 1, \dots , m$	$\sum_{i=1}^{n} x_{ij} \, \lambda_j \leq x_{io} \, , i = 1, \dots , m$
$\sum_{i=1}^{n} y_{rj} \, \lambda_j \geq y_{ro} \, , r = 1, \dots , s$	$\sum_{i=1}^{n} y_{rj} \, \lambda_j \geq \theta y_{ro} \, , r = 1, \dots , s$
$\lambda_j \geq 0 \, , j = 1, \dots , n$	$\lambda_j \geq 0 \, , j = 1, \dots , n$

Figure 17 - DEA Dual Model formulation for CRS methodology (Cooper et al., 2006)

4.1.1.2 DEA – VRS Model

The BCC models appeared to be necessary to include the possibility of the existence of VRS technologies on CRS formulations, either increasing or decreasing. (Banker et al., 1984) On this new model, efficiency is broken down into technical efficiency (TE) and scale efficiency (SE), which did not occur on CRS models, where inefficiency may be due either to inefficiency in the production process or the use of a non-optimum scale. On the VRS model, this distinction is made based on the addition of a constraint that ensures a comparison of the DU being analysed with the convex combination of DUs of the sample (R. Marques & Silva, 2006). This constraint demands that a DU can only be compared with a DU of the equivalent size. On the VRS models, it is notable that the efficiency is always greater than or equal to the previous CRS method.

In figure 18, it is possible to see the DEA-BCC model, where the points A, C and F are considered efficient.

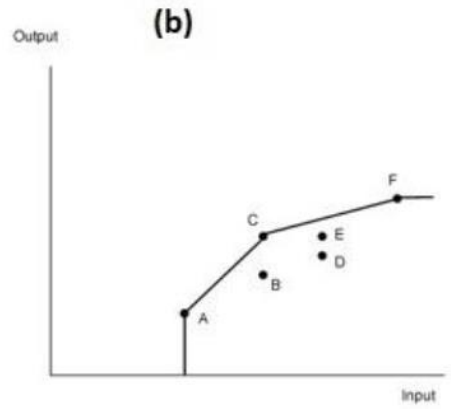

Figure 18 - DEA-BCC model example of production frontier (Kim & Harris, 2008)

The slight difference between the two methods, CRS and VRS, is related to the addition of the convexity constraint given by the expression $\sum k\, \lambda k = 1$. This can be seen in figure 19 as an illustrative example of the primal VRS model.

Primal Model

DEA – BCC input-oriented	DEA – BCC input-oriented
$\text{Max } z = \sum_{r=1}^{s} u_r\, y_{ro} - u_0$	$\text{Min } z = \sum_{i=1}^{m} v_i\, x_{io} - v_0$
With the constraints:	With the constraints:
$\sum_{i=1}^{m} v_i\, x_{io} = 1$	$\sum_{r=1}^{s} u_r\, y_{ro} = 1$
$\sum_{r=1}^{s} u_r\, y_{rj} - \sum_{i=1}^{m} v_i\, x_{ij} - u_0 \leq 0$	$\sum_{i=1}^{m} v_i\, x_{ij} - \sum_{r=1}^{s} u_r\, y_{rj} \leq 0$

Figure 19 - DEA Primal Model formulation for VRS methodology (Cooper et al., 2006)

Consequently, from this first formulation, a dual formulation can be done as exposed in figure 20.

Dual Model

DEA – BCC input-oriented	DEA – BCC input-oriented
$Min\ \theta$	$Max\ h_0$
With the constraints:	With the constraints:
$\theta x_{io} - \sum_{i=1}^{n} x_{ij}\, \lambda_j \geq 0$	$h_0 y_{jo} - \sum_{i=1}^{n} y_{rj}\, \lambda_j \leq 0$
$-y_{ro} + \sum_{i=1}^{n} y_{rj}\, \lambda_j \geq 0$	$-x_{ro} + \sum_{i=1}^{n} x_{rj}\, \lambda_j \geq 0$
$\sum_{j=1}^{n} \lambda_j = 1$	$\sum_{j=1}^{n} \lambda_j = 1$
$\lambda_j \geq 0\,, \forall\, k$	$\lambda_j \geq 0$

Figure 20 - DEA Dual Model formulation for VRS methodology (Cooper et al., 2006)

4.1.2 Multiple Linear Regression

The linear regression method is a prevalent mathematical method related to a dependent variable named Y and one or more independent X variables. The purpose of the approach is to predict, verify and evaluate how a change in the conditions that interact with the variable Y affects

the predicted output variable, Y. In addition to X variables, there are also parameters generated from the data included in the model. These parameters are identified by ***β***.

The linear regression model can be divided into simple linear regression and multiple linear regression. The difference is in the number of independent variables used to determine the expected behaviour of *Y*. The linear relation between *Y* and *X* is defined by:

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_k x_k + \varepsilon_i \quad (1)$$

With each parameter meaning:

- Y: dependent variable

- β_0: fixed technical coefficient, which is the minimum (or maximum, depending on the slope of the regression) value of *Y*

- β_j, *j* = 1, …, k: technical coefficients related to each independent variable

    - Represents the impact each unit of *X* takes on the value of *Y*

- x_j: independent variables

- ε_i: the deviation between each actual *Y* observation and the predicted value of *Y* generated by the model

When performing an MLR, some assumptions should be considered:

1. The regression residuals should be shown as a normal distribution
2. MLR considers that the multicollinearity in the data is inexistent – indicates when independent variables are highly correlated to one another
3. MLR also assumes homoscedasticity, which means that across the values of independent values, the variance of error terms is approximately the same (Statistics Solution, 2020)

4.1.2.1 Least-Square Method

This method consists of defining a hyperplane in k-dimensional space to fit better into the data in use. The focus is on residuals, which are the difference between observed and predicted values of the dependent variable. This method allows minimizing the sum of the residuals' square (Marill, 2004). Considering $\hat{Y}_t$ as the predicted value for the dependent value through the model, and Y_i as the observed value, the residual of each observation (e_i) of the model is taken by:

$$e_i = \hat{Y}_t - Y_i \quad (2)$$

Regarding the Sum of Squares Error/Residuals and being n the number of observations, which is the squared of each residual on the observations is taken by:

$$SS_E = \sum_{i=1}^{n} {e_i}^2 \quad (3)$$

To summarize, and as mentioned before, the goal of the Least-Square Method is to minimize SS_E (Sum of Squared Errors). The regression adjustment is as better as more minor is the SS_E since the difference in the predicted and observed value is smaller.

4.1.2.2 Model's significance (F-test)

This test included in the linear regression model is used to verify that every independent variable X in the model adds significantly to data to describe the variance of the dependent variable linearly. The F-test is performed with the requirement of the error term (ε_i) being normally distributed and independently with a mean equal to 0. The hypothesis test to perform is taken by:

$$H_0: \beta_1 = \beta_2 = \ldots = \beta_k = 0 \text{ (null hypothesis)} \quad (4)$$

$$H_1: \beta_j \neq 0 \text{ (alternative hypothesis)} \quad (5)$$

These hypotheses are, as mentioned earlier, related to having independent variables X in the model capable of explaining the variation of the dependent variable linearly. The null hypothesis (4) assumes that the model can predict values for the dependent variable. If this hypothesis is rejected, there is at least one variable capable, and then the alternative hypothesis is accepted.

The test is based on the equation of Sum of Squares (3), defined above, and the following ones with $\overline{Y_i}$ being the mean of the observed output variables:

$$SS_T = \sum_{i=1}^{n} (Y_i - \overline{Y_i})^2 = SS_R + SS_E \quad (6)$$

Taken that:

$$SS_T = \sum_{i=1}^{n} (\hat{Y}_t - \overline{Y_i})^2 \quad (7)$$

The main factor of the whole F-test is the F_0 since this is the value that defines whether the null hypothesis is rejected or accepted. F_0 is given by:

$$F_0 = \frac{\frac{SS_R}{k}}{\frac{SS_E}{(n-k-1)}} \quad (8)$$

The rejection criterion is given by:

$$F_0 > f_t = f_{\propto}[k, n-p] \quad (9)$$

The null hypothesis is rejected, with a significance level α, if F_0 is bigger than f_t – which is the value established for F distribution – with *k* degrees of freedom, and (*n* – *k*) residuals' degrees of freedom (dF). If the null is rejected, it is concluded that the model is significant with a (1 – α) x 100% confidence. Moreover, at least one independent variable is a good predictor for the dependent variable. These interpretations can be made by analysing the ANOVA table on the output. If the p-value given is lower than α, then the null hypotheses (H_0) is rejected, and, consequently, the model is significant. On the other hand, one important aspect must be considered: the model being considered significant is not the most relevant result of predicting the outcome variable (Bremer, 2012).

4.1.2.3 Coefficient of Determination (R2)

One of the most used methods to analyse the model's suitability is the coefficient of determination, more specifically measuring the fitting of the model, which is the R^2 – accessing how accurate the model fits. This coefficient concluded regarding the proportion of the dependent variable that can be explained by all variables which predicted its value (Renaud & Victoria-Feser, 2010). The following expression gives R2:

$$R^2 = \frac{SS_R}{SS_T} \quad (10)$$

This coefficient is expressed with values from 0 to 1. As R^2 gets closer to 1, more remarkable is the proportion of the dependent variable explained by the independent variables. On the other hand, a high value of this coefficient guarantees that the regression model is accurate for the data in question. When a variable is added, the coefficient value always increases (adding a variable also increases the Sum of the Squares). Consequently, it is given that adding new variables into the model will not contribute to the efficiency of the model. Thus, models with a high coefficient of determination can produce more unreliable predictions (Schneider et al., 2010).

Some researchers studied for a solution to this issue and found that, for these cases where the model has more than one independent variable, it should be used the adjusted coefficient of determination – defined by R^2_{Adj}. This new coefficient is calculated using:

$$R^2_{Adj} = \frac{\frac{SS_E}{(n-p)}}{\frac{SS_T}{(n-1)}} = 1 - (1 - R^2)\left(\frac{n-1}{n-p}\right) \quad (11)$$

The difference is that this last new coefficient considers the number of independent variables. Consequently, if a new input variable is included in the model, the coefficient value

generally decreases. Thus, if there is a considerable difference between R^2_{Adj} and R^2, non-significant variables are most likely included in the model. To summarize, R^2_{Adj} helps to penalize the insertion of non-essential variables when calculating the relation with the dependent variable (Bremer, 2012).

4.1.2.4 Variable's significance

As mentioned before, the importance of checking individually each variable included in the regression is crucial to search for the potential of each one. If, accordingly to the F-test, the model is significant, there is at least one variable in the model capable of predicting the output variable. Thus, each variable is essential to minimize the weakness of the degree of adjustment when added a non-explanatory variable (Siegel, 2016).

This analysis is made by analysing the p-value of each variable. If the value is smaller than the significance degree α, then the variable in question is significant in the model. However, if the value is higher than the significance degree α, the variable in question is not relevant to the model and can be excluded from the model (Bremer, 2012). The p-value is related to testing the null hypothesis that each variable is not correlated with the dependent variable. If this is confirmed and there is no correlation between independent and dependent variables, the p-value is higher than the significance degree α, and the null hypothesis is accepted.

On the other hand, as mentioned earlier, if the p-value is lower than α, the null hypothesis is rejected. The variable is assumed as an explanatory variable of the dependent one (Frost, 2018).

4.1.2.5 Model with Interaction Terms

On the other topics, it was seen that one or more independent variables explain the dependent variable of the models. However, an independent variable can be dependent as well on the magnitude of another explanatory variable. For instance:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_1 . x_2 \quad (12)$$

Consequently, the effect of x_1 on the independent variable is given by:

$$\frac{\Delta y}{\Delta x_1} = \beta_1 + \beta_3 x_2 \quad (13)$$

From (13), it is concluded that if $\beta_3 > 0$, the impact of one more unit in x_1 affects the dependent variable more when x_2 also increases. This is called the interaction effect between two variables. When a model has interaction terms, it becomes essential to distinguish which parameters are accurate to analyse. For instance, β_2 is the effect of x_2 on the dependent variable

when $x_1 = 0$. On variables that interact with each other, this parameter will most likely not be interesting to analyse (Wooldridge, 2012).

5. Methodology Application

This chapter aims to explain and show which data sets were used and how they were collected, the process of choosing which variables to include in the models, and how the models were executed. Thus the chapter is divided into three subtopics: data collection, variables and model.

5.1 Data Collection

The data collection started with deciding which leagues would take part in the study. The big-5 – Spain, Germany, England, Italy, France – were predictable since they are the leagues with the most impact worldwide. This impact is economically speaking and in terms of popularity, as it was possible to observe in the second chapter of this dissertation. Moreover, these leagues are the ones who contribute the most with winners of the UEFA Champions League and the FIFA Club World Cup, which are the most important clubs' competitions in the world. The insertion of the Portuguese football league – Liga NOS – is also easily explained by the assiduous presence of 2 or 3 clubs in the Champions League. Besides this, Portugal is getting close to France on the UEFA ranking, and the overall club comparison is similar – except for Paris Saint-Germain (UEFA, 2020).

Secondly, the step was to decide in which seasons the study would focus. The intention was to consider the most recent seasons and the most recent data. However, one detail impacted the decision: Ligue 1 2019/2020 (France) was not concluded due to the pandemic situation that occurred in 2020. Consequently, the data from the rest of the leagues would be disproportional in comparison to France. Thus, the decision was to collect data from 4 seasons to have three variations from one year to another to include in the study.

- 1st season of the study: 2015/2016
- 2nd season of the study: 2016/2017
- 3rd season of the study: 2017/2018
- 4th season of the study: 2018/2019

The dataset collection started with aggregating all the clubs that participated in these leagues, divided by league and year. With these lists, it was possible to understand which clubs were qualified and disqualified from year to year and will not have enough data to be included in the model. Then, the minutes of all players participating in these leagues by the club were collected from the website and data store ZeroZero – each data set collected was of all players fielded from each club on each season (ZeroZero, 2020). This data would then be used to calculate some variables such as the percentage of new players besides the same percentage although from the players that already were in the club the previous season. Afterwards, the data related to the clubs were collected. This data was collected from SoccerStats and TransferMarkt,

two major websites and data stores related to football worldwide (SoccerStats, 2020; TransferMarkt, 2020a). This data compilation resulted in many variables:

- Games Played
- Points
- Victories
- Draws
- Losses
- Goals Scored
- Goals Conceded
- Goal Average (Difference between scored and conceded)
- % Clean Sheets (games percentage without conceding a goal)
- % Failed to Score (games percentage without a goal scored)
- Players Fielded
- Squad Value

Additionally, on 2016/2017, on 2017/2018 and 2018/2019, since all had a previous season included in the study, it was also calculated four additional variables regarding variations:

- Sum of New Players Minutes
- Sum of Previous Season Players Minutes
- % of New Players Minutes
- % of Previous Season Players Minutes

Finally, all variables were calculated to establish the variation from season to season. After the aggregation of both collected and calculated variables, the matrix of all variables included 464 observations on the totality of the clubs.

5.2 Variables

In chapter 3, it was possible to analyse multiple studies regarding performance and turnover to collect some insights into interesting variables to include in this dissertation. The studies included variables like the salary turnover, previous year's attendance, current and previous season winning percentage, number of championships, city population, recently built arena, expenses on transfers and contract renewals, operational costs, total expenses, profit

margin, assets to debt ratio, club's age, location or presence on international competitions. These variables are from different areas of investigation such as finances, demographs or performance; however, all can be important for explaining player turnover. In terms of output variables, the studies were more similar and less divergent in terms of using different variables. The variables used were points earned, goal average and total attendance. Of course, these variables represented all variables used in studies compared with this dissertation and were not all included.

First, it was done a covariance table to help choose which variables to use. This table aims to analyse which variables can be related to each other and minimize the impact of one variable hiding the impact of the other. Thus, this table confirmed that some variable sets were not possible together in a model. It was assumed that any correlation value (in absolute value) above 0.6 was too high and could interfere with the model results. The essential variable and the first chosen was one related to turnover directly. The options were: Sum of New Players Minutes, Sum of Previous Season Players Minutes, percentage of New Players Minutes and percentage of Previous Season Players Minutes. In order to be different from the previous studies and to have a more consistent and more accessible way to analyse, the primary variable chosen was the percentage of New Players Minutes. This variable is critical since just the variable itself can give essential information regarding how important the goal of this dissertation is and highlight the balance of new players and "old" players on every squad from 2016 to 2019 in the six major leagues.

On the other hand, the usage of this variable requires information from two seasons, which results in the significant limitation of this dissertation. Clubs that were present in the league in the previous season are not considered. Consequently, the observations are cut from 464 to 302 in total.

The process of choosing the rest of the input variables was complicated. However, some were highlighted and tested after to conclude regarding the insertion in the model. The list of possibilities was:

- Squad Value: essential to distinguish a big club from a small club
- Goals Scored: essential to analyse the offensive performance of the team
- Goals Conceded: essential to analyse the defensive performance of the team
- % Failed to Score: also related with the offensive performance, however on a different approach
- Players Fielded: number of players the club used for the entire season

The choice was squad value, and the number of players fielded. The first is because, as mentioned above, is vital to distinguish between big and small clubs. The covariance values with the variables involved are all above 0.2 and below 0.6 (in absolute value), which is a variable related to the remaining variables. The second variable, because the first has significant

correlation values and has a negative correlation, is vital to increase the understatement of results. Both Goals Scored or Goals Conceded have either values out of the range or values close to zero on the covariance, so it was decided to take them out of the DEA model; however, it is included in the regression control variable.

Regarding the output variables, there were some options analysed, such as:

- Points: points earned from victories and draws
- Variation of points: difference of points from the current season to the previous one
- Goals Scored
- Goals conceded
- Goal Average: the difference between scored and conceded goals

Both goals scored and conceded were easily eliminated since they did not have robust results that could justify what was expected. Also, goals scored when performed alone, and with %NPMin showed results with a high number of clubs with maximum efficiency and no evidence was concluded from those results. The variable of goal average proved to have results very similar to the variable points. Also, this can be justified since both have numbers of covariance very similar with all the input variables. Note that this variable was one of the most used in studies related to football. In this dissertation, the variable “goal average” is no more than confirming what is observable with the output “points”. The output variable points will be used better to control the results regarding the variation of points. Thus, the two output variables used were points and variation of points. The second one is important to analyse the impact of the new players on the difference in points from the previous season, which is the main focus of the study.

In conclusion, on the input side, the variables are the percentage of minutes of new players (%NPMin), squad value (SqValue), and the number of players fielded (PFielded). On the output side, points (Pts) and variation of points (VarPts) were the choices as mentioned before. Next, it will show some statistics regarding the variables to understand better which values the variables take.

Table 1 - Statistics from used dataset

	Description	Average	MinValue	MaxValue
%NPMin	Percentage of New Players Minutes	34,18%	2,95%	79,87%
SqValue	Squad Value. Essential to distinguish a big club from a small club	201,94M €	11,48M €	1160M €
Pfielded	Number of players the club used for the entire season	27,71	21	41
Pts	Points earned from victories and draws	52,52	16	100
VarPts	Difference of points from the current season to the previous one	-1,50	-44	43

- **% New Player Minutes:**
    - The average from all clubs for the six leagues was 34,18%
    - The minimum value found was on Premier League with 2.95% (Tottenham Hotspur FC – season 2018/ 2019)
    - The maximum value found was on Liga NOS – which has the highest values on every season and with values almost 10% higher – with 79,87% (Moreirense FC – season 2017/2018)
    - As expected, this is a variable with a significant deviation

- **Squad Value**
    - The average from all clubs for the six leagues was 201,94 Million Euros
    - The lowest value was observed on Liga NOS – justifiable by the smaller economic size in comparison with the rest – with 11,48 Million Euros (CD Aves – season 2018/2019)
    - The highest value was observed on La Liga with 1160 Million Euros (FC Barcelona – season 2018/2019)
    - Also, as expected, and justified by the growth of big clubs in modern football, the deviation from big club to small club is quite noticeable

- **Players Fielded**
    - The average from all clubs for the six leagues was 27,71 players
    - There were three clubs with the same value, which was the minimum observed, all on 2018-2019 season, with 21 players fielded - Bayer 04 Leverkusen (Bundesliga), Brighton & Hove Albion FC and Manchester City FC (Premier League)
    - The highest value was observed on Ligue 1 with 41 players fielded (AS Monaco FC – season 2017/2018)

- **Points**
    - The average points from all clubs for the six leagues were 52,52
    - The lowest score in points observed occurred on Premier League with 16 (Huddersfield Town AFC – season 2018/2019)
    - The highest score in points observed also occurred on Premier League with 100 (Manchester City FC – season 2017/2018)
    - Note that Bundesliga and Liga NOS have fewer points since both have only 18 teams instead of 20 as the other four leagues. Consequently, each team play less than four games. Even though this is not a variable with a significant deviation

- **Variation of Points**
    - The average of the difference in points of one season from the previous one was -1,50 points
    - The lower variation was recorded on Ligue 1 with -44 points (AS Monaco FC – season 2018/2019)
    - The highest variation occurred in Premier League with 43 points (Chelsea FC – season 2016/2017)
    - As expected, this variable has a more significant deviation than the variable “Points” and interestingly has more observations with negative than with positive variation

The main variables of interest are the input side, the percentage of minutes of new players, and the variation of points on the output side. Both variables are related to the change from one season to the next, and the goal of the dissertation is indeed analysing the impact of that change.

5.3 Model

The methodology applied was two different approaches to have more robust results and analyse the data differently. First, was performed a DEA with the data and secondly, a multiple linear regression.

DEA-VRS (also called DEA-BCC) was used, which was fully detailed in chapter 4, with an input orientation. The software allowed to have the results was MATLAB, and the toolbox was retrieved from Álvarez et al. (2016). The choice of this model came after analysing the potential of both DEA-CRS and DEA-VRS. The first was the first DEA model performed by Charnes, Cooper and Rhodes. Generally, DEA-CRS gives more conservative results since it does not divide the efficiency into Technical Efficiency and Scale Efficiency. Also, this was proved by testing the model on the MATLAB DEA toolbox that was used. The results were difficult to analyse when the model performed was DEA-CRS since the variety of efficiency scores was lower. Also, many DMUs were scored with 0 on the efficiency ranking. From this evidence, it could be concluded that DEA-CRS would be more effective with a model with a higher number of independent variables. However, this was refuted since the number of zeros was extremely high on all the observations. Consequently, the model chosen to be used was DEA-VRS, even with the downside of sometimes having an excess of maximum efficiency DMUs.

First, the DEA model was performed on each league and each year. The main reason was to avoid club repetitions and comparisons between teams that do not compete against each other. Thus, each output of the model is related to a specific league season. The goal was to conclude regarding the impact of the new players on the performance of the team. For that reason, the primary analysis was done with %NPMin as the only independent variable and VarPts as the dependent variable. Besides this, it was also performed DEA analysis with a combination of the other two variables – SqValue and PFielded – with the primary independent variable – %NPMin. These combinations were done both with VarPts and Pts as the dependent variable. In figure 21, the steps until the conclusions of the DEA model are shown.

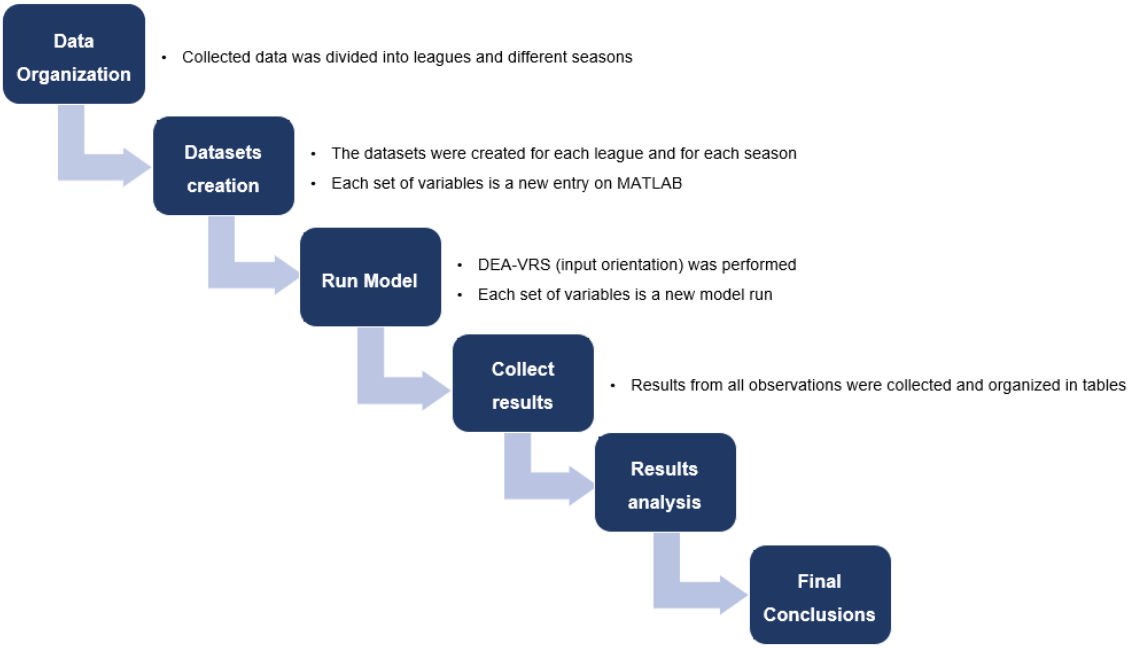

Figure 21 - DEA-VRS steps

Secondly, to compare the results of the DEA model and analyse whether the input variables can predict the behaviour of the dependent variable, a linear regression was performed in STATA software. Each model was performed with fixed effects on league and year, i.e. the model is performed without affecting the league and the year, which allows avoiding analysing the same clubs and comparing clubs from different leagues (which do not compete against each other). The linear regression also allowed the understanding of whether exist any variable interaction. This interaction happens when a variable can only take different values if another variable also changes to achieve the output.

6. Results

The current chapter presents the main results and interpretations that may be drawn from the obtained results. The purpose is to answer the main dissertation question: “Does the squad stability affect performance and team efficiency?”. From the studies analysed in chapter 3 and general opinion, the answer is easily said to be a positive one for all the reasons explored by the authors exposed in that chapter. Thus, subsequently, it is shown whether the results from this study confirm or refute those.

Across the results, it is easily noticeable that there is no direct or extreme relationship between a low percentage of new players minutes and efficiency (when the output variable is VarPts). Instead, what is noticeable is that generally, a lower percentage is related to higher efficiency, although with some outliers that can be explained with the behaviour regarding VarPts.

6.1 DEA results

In table 2, it is possible to observe the DEA-VRS results for the 2018-2019 English Premier League season (notice that, as mentioned earlier, not all clubs were considered due to these not being on the league on the previous year, since those cases did not have enough information for this model). The two lowest results of the study’s input variable were given the maximum efficiency, 1 – Manchester City FC and Tottenham Hotspur FC – even with negative values of the output variable. On the other hand, Liverpool FC, with approximately one new player in every five in the squad – a seven times higher proportion than the previous two clubs – also had the maximum efficiency value. However, note that the variation of points on the Liverpool case is notable since they made the most significant recovery from the previous season. Consequently, the efficiency rate is easily justified by that result. Another interesting case is West Ham United FC. They had almost half of the squad made up of new players, even though they had a rise of 10 points compared with the previous season – the second-highest score on this EPL season. However, West Ham is in the centre of the table regarding efficiency. A similar %NPMin of Liverpool is Brighton & Hove Albion FC, although their performance compared with the previous season declined. So those new players did not correspond to a reasonable efficiency rate.

Regarding the cases with higher %NPMin, they are not directly connected to a lower value of efficiency, as seen in the case of West Ham United FC, previously presented. Moreover, Chelsea FC, the club with the second-highest %NPMin, is the third-lowest score inefficiency. Additionally, if Manchester City, Tottenham (the two lowest), West Ham and Chelsea are not considered, the values of %NPMin are relatively uniform, around 20-30%. Consequently, the efficiency values are more influenced by the performance regarding the variation of points than regarding the independent variable. Huddersfield Town AFC, Manchester United FC and Burnley FC, which are the three lowest in the variation of points, are all below the middle of the table when sorted by efficiency.

Table 2 - DEA-VRS results of EPL 2019 with %NPMin and VarPts

DMU	%NPMin (X)	VarPts (Y)	Efficiency
Liverpool FC	21,0%	22	1
Manchester City FC	3,6%	-2	1
Tottenham Hotspur FC	3,0%	-6	1
Watford FC	23,2%	9	0,4967
Crystal Palace FC	19,3%	5	0,4484
Southampton FC	23,8%	3	0,3016
Arsenal FC	33,8%	7	0,2982
Everton FC	30,1%	5	0,2872
West Ham United FC	45,9%	10	0,267
Leicester City FC	32,7%	5	0,2639
Newcastle United FC	24,9%	1	0,23
Burnley FC	13,3%	-14	0,2222
AFC Bournemouth	26,8%	1	0,2139
Manchester United FC	14,0%	-15	0,2112
Chelsea FC	39,0%	2	0,1658
Brighton & Hove Albion FC	20,0%	-4	0,1624
Huddersfield Town AFC	21,1%	-21	0,14

Assuming that Liverpool was an outlier in this dataset, the model was performed again without that DMU. In table 3, there are results of that model. Manchester City and Tottenham had the same result, which was expected accordingly with their lower percentage of new player minutes. On the other hand, Watford FC and West Ham United, even with a higher %NPMin, have the two best VarPts and were scored with the maximum value of efficiency as Liverpool. Thus, this model proved the stated above when Liverpool was included, and the conclusion is that Liverpool was not such an outlier. Teams with the best recovery in points in comparison to the previous season have the maximum efficiency. Except for Manchester City and Tottenham, the clubs are almost ordered from the highest to the lowest variation of points in this case. Note that the exceptions of this ordering are due to a higher %NPMin, which also penalizes the team – such as the cases of Brighton & Hove Albion FC and Chelsea FC. This rank can be justified from what was mentioned above in the case with all the teams. When the independent variable results become more homogenous, the dependent variable influences the results (VarPts).

Table 3 - DEA-VRS results of EPL 2019 with %NPMin (Without Liverpool)

DMU	%NPMin	VarPts	Efficiency
Manchester City FC	3,6%	-2	1
Tottenham Hotspur FC	3,0%	-6	1
Watford FC	23,2%	9	1
West Ham United FC	45,9%	10	1
Crystal Palace FC	19,3%	5	0,8346
Arsenal FC	33,8%	7	0,5808
Everton FC	30,1%	5	0,5345
Southampton FC	23,8%	3	0,5246
Leicester City FC	32,7%	5	0,4912
Newcastle United FC	24,9%	1	0,3578
AFC Bournemouth	26,8%	1	0,3327
Chelsea FC	39,0%	2	0,2749
Burnley FC	13,3%	-14	0,2222
Manchester United FC	14,0%	-15	0,2112
Brighton & Hove Albion FC	20,0%	-4	0,1624
Huddersfield Town AFC	21,1%	-21	0,14

In order to have a better overview regarding the efficiency rates and the relation with %NPMin, in table 4, there is a summary of the results with %NPMin as the input and VarPts as the output on all the six leagues on 2018-2019 season. The rest of the leagues have similar results regarding what has been explained above in the English Premier League. The lowest value on %NPMin has, in all leagues, the maximum value of efficiency, even if the performance has declined compared with the previous season, evident in Ligue 1, La Liga and Liga NOS – which was also observable in Manchester City and Tottenham on the above analysis. There are also cases compared with the Liverpool case, and even with more noticeable percentages – Ligue 1 and Liga NOS cases, with 48,9% on %NPMin and +37 on VarPts and 64,4%, +20, respectively.

Moreover, in all leagues in every season, there are similar cases with the related above. Both 2018 and 2017 have similar results as 2019. It is important to highlight that:

- The team with the lowest %NPMin is permanently assigned with maximum efficiency
- The team with the highest variation of points regarding the previous season is permanently assigned with maximum efficiency

Table 4 - DEA-VRS results for all leagues in 2019

Premier League 2019			Bundesliga 2019			Ligue 1 2019			La Liga 2019			Liga NOS 2019			Serie A 2019		
%NPMin	VarPts	Effic	%NPMin	VarPts	Effic	%NPMin	VarPts	Effic	%NPMin	VarPts	Effic	%NPMin	VarPts	Effic	%NPMin	VarPts	Effic
21,0%	22	1	10,8%	0	1	13,6%	-16	1	17,9%	-4	1	16,2%	-3	1	20,0%	9	1
3,6%	-2	1	11,4%	13	1	25,6%	5	1	18,2%	-2	1	30,4%	6	1	20,8%	4	0,96
3,0%	-6	1	23,5%	22	1	48,9%	37	1	21,0%	4	1	64,4%	20	1	23,1%	2	0,86
23,2%	9	0,50	13,2%	-6	0,82	31,5%	11	0,95	25,2%	10	1	30,1%	1	0,75	23,1%	-13	0,86
19,3%	5	0,45	15,4%	-9	0,70	26,4%	2	0,90	19,1%	-8	0,94	55,8%	9	0,68	24,3%	-12	0,82
23,8%	3	0,30	15,9%	3	0,69	25,9%	-2	0,83	21,1%	-6	0,85	38,1%	2	0,63	31,6%	-5	0,63
33,8%	7	0,30	38,0%	21	0,58	23,5%	-6	0,82	21,3%	-14	0,84	52,6%	6	0,58	32,4%	-23	0,62
30,1%	5	0,29	19,5%	8	0,58	34,8%	8	0,80	24,1%	-3	0,75	30,6%	-8	0,53	34,4%	-3	0,58
45,9%	10	0,27	24,7%	11	0,46	18,6%	-14	0,79	26,8%	1	0,73	31,5%	-3	0,51	34,4%	-16	0,58
32,7%	5	0,26	24,3%	-4	0,45	28,8%	-4	0,71	32,4%	4	0,65	32,3%	-8	0,50	35,0%	9	0,57
24,9%	1	0,23	30,4%	0	0,36	34,7%	-6	0,56	27,8%	-12	0,64	33,6%	-11	0,48	42,2%	-1	0,47
13,3%	-14	0,22	33,7%	-30	0,32	29,2%	-14	0,51	34,7%	3	0,59	71,6%	4	0,38	43,3%	-11	0,46
26,8%	1	0,21	34,8%	-18	0,31	48,0%	1	0,48	38,0%	1	0,52	45,4%	-4	0,36	43,4%	4	0,46
14,0%	-15	0,21	35,3%	-23	0,31	64,1%	11	0,47	40,4%	-8	0,44	45,9%	-15	0,35	47,3%	0	0,42
39,0%	2	0,17	36,5%	5	0,30	30,6%	-20	0,45	40,4%	-10	0,44	56,6%	-1	0,34	51,5%	5	0,39
20,0%	-4	0,16	42,1%	7	0,27	49,1%	-5	0,40	50,9%	-17	0,35	69,5%	-6	0,23	57,5%	-3	0,35
21,1%	-21	0,14	-	-	-	54,9%	-7	0,34	71,1%	2	0,28	-	-	-	59,4%	3	0,34
-	-	-	-	-	-	47,3%	-44	0,29	-	-	-	-	-	-	-	-	-

After analysing table 4, and with those results, a graph regarding the tendency of the relation between %NPMin and Efficiency was essential to have a better view of how the variables behave. Thus, in figures 22, 23 and 24, the graphs show the tendency line of the results on all the leagues of the study per year.

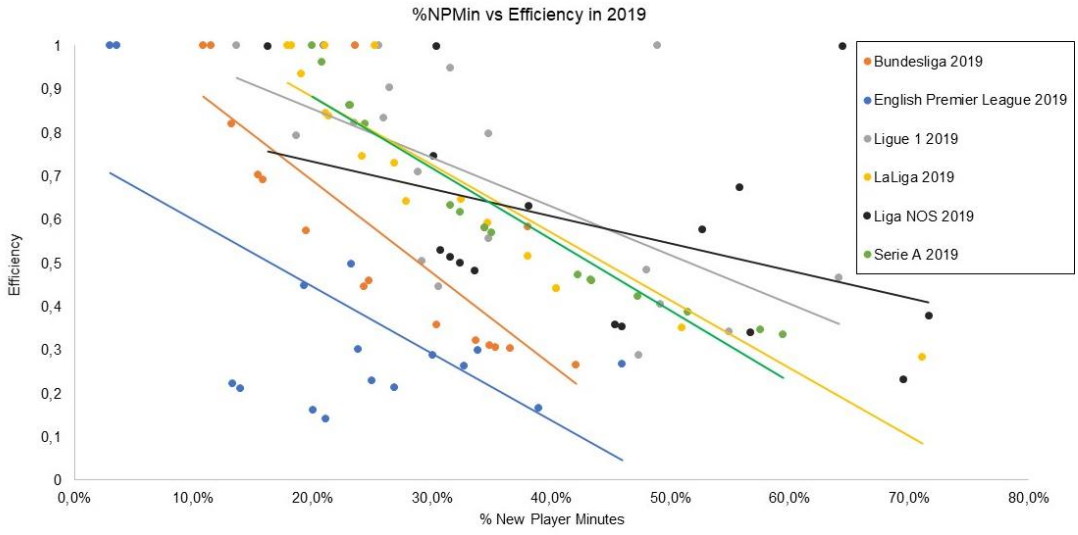

Figure 22 - Tendency Line on %NPMin vs Efficiency 2019

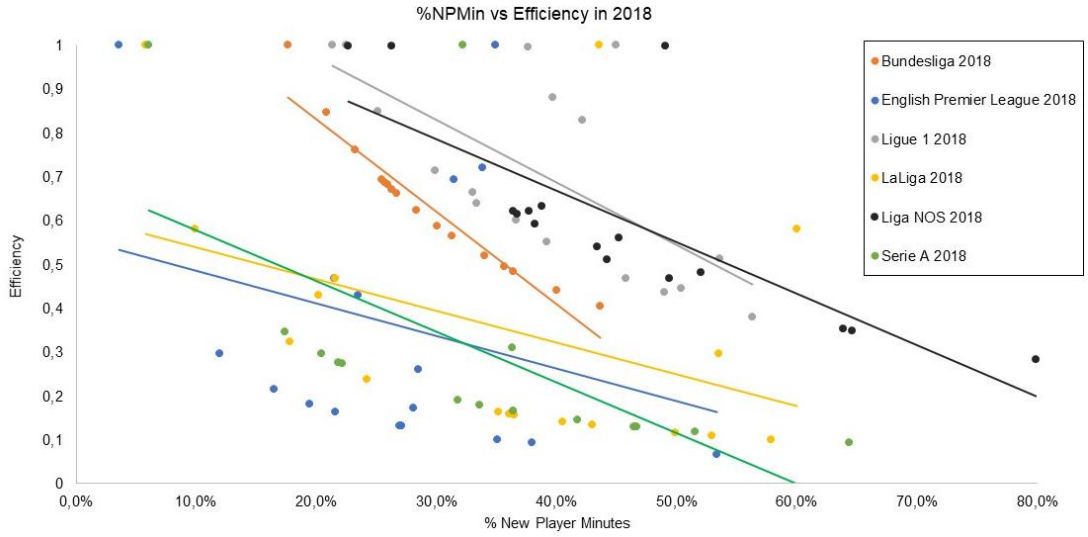

Figure 23 - Tendency Line on %NPMin vs Efficiency 2018

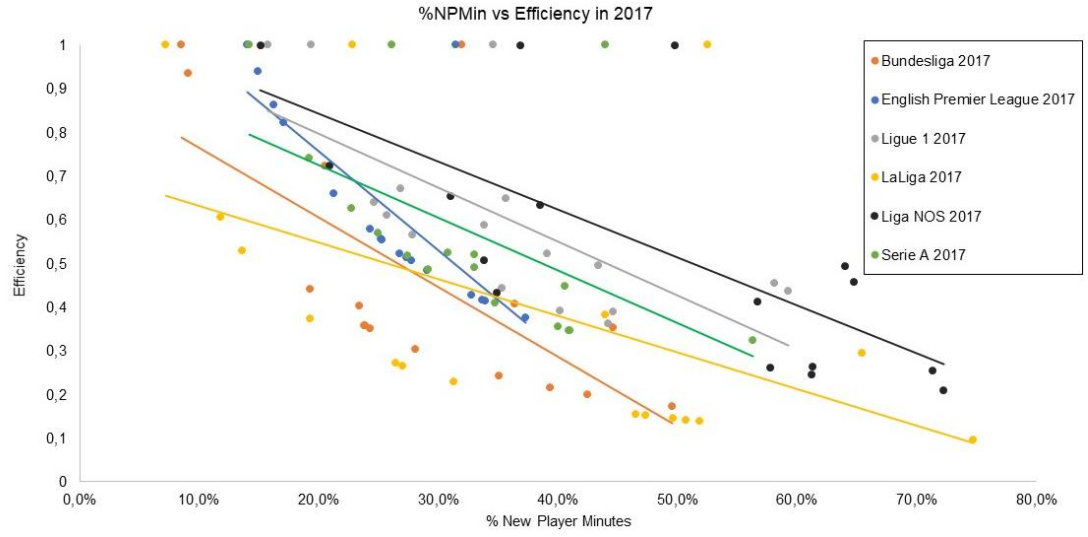

Figure 24 - Tendency Line on %NPMin vs Efficiency 2017

From these figures (22, 23 and 24), it can be observed that every year on every league the tendency is to relate significant percentages of new players with lower efficiency rates – in other words, the more stable the squad from one year to another the better the efficiency. Also, it is important to note that the league where this is more perceptible over the three years is the Bundesliga, only surpassed by the Premier League in 2017. In conclusion, all seasons proved that when the team's stability decreases – corresponding to a higher %NPMin – the value of efficiency also decreases.

One of the study's limitations, which can be observed in table 5, occurs when the model is performed with more than one input variable. Some datasets with two variables could reveal some meaningful information, where it was possible to extract some conclusions. In table 5 it can be observe that when the model is performed with all the variables, the results become impossible to analyse and conclude. Ten out of sixteen clubs were assigned with maximum efficiency. Although, the majority presented several DMUs with maximum efficiency, making it impossible to draw any conclusions from them.

Table 5 - DEA-VRS results of Bundesliga 2019 with all variables

DMU	%NPMin	SqValue	PFielded	VarPts	Efficiency
1. FSV Mainz 05	0,4209	86,6	28	7	1
Bayer 04 Leverkusen	0,1585	359,35	21	3	1
Borussia Dortmund	0,3799	376,8	23	21	1
Borussia Monchengladbach	0,1948	191,53	22	8	1
FC Augsburg	0,154	100,95	30	-9	1
Hertha BSC	0,1083	125,18	26	0	1
RB Leipzig	0,1144	314,45	23	13	1
SC Freiburg	0,3036	94,28	28	0	1
SV Werder Bremen	0,2472	110,85	24	11	1
VfL Wolfsburg	0,2354	145,53	25	22	1
FC Bayern München	0,1319	835,55	24	-6	0,9402
Hannover 96	0,3482	98,28	32	-18	0,9336
VfB Stuttgart	0,353	159,45	26	-23	0,893
TSG 1899 Hoffenheim	0,2429	188,63	29	-4	0,8073
Eintracht Frankfurt	0,3654	139,55	30	5	0,7994
FC Schalke 04	0,3367	243,9	32	-30	0,703

In table 5, there is not an evident pattern of variable set behaviour that leads to a maximum efficiency score. Figures 22, 23 and 24, show the influence of a low %NPMin on efficiency and a negative correlation between an increase in %NPMin and efficiency. Besides %NPMin, only the dependent variable VarPts shows a noticeable tendency on influencing the efficiency score – generally, a higher variation of points results in higher efficiency. The remaining variables show minimal tendency on influencing efficiency. However, generally, the behaviour is:

- **SqValue:** lower SqValue reflects on a higher efficiency
- **PFielded:** lower PFielded reflects on a higher efficiency

The DMUs assigned with maximum efficiency show at least one variable influencing the most the efficiency score. On the other hand, there is not an ideal combination of variables that reflect maximum efficiency. Thus, it is acceptable that the result, when combining all variables, shows a high number of DMUs with a maximum efficiency score. Later with the results of linear regression, it will be possible to conclude the weight of each variable contributing to the variation

of points. Subsequently, the results with %NPMin, together with SqValue, and together with PFielded, will be analysed.

Table 6 shows the DEA results for the model with %NPMin and SqValue on Bundesliga 2018-2019. As mentioned earlier, the limitation of a high number of maximum efficiency scores when the number of independent variables is more than one, is also evident in this case. The first observations from the table are:

- The two clubs with the lowest %NPMin have efficiencies equal to 1 (Hertha BSC and RB Leipzig)
- The club with the lowest SqValue has also an efficiency equal to 1 (1. FSV Mainz 05)
- The club with the highest VarPts has an efficiency equal to 1 (VfL Wolfsburg)
- The club with the lowest VarPts also has the lowest efficiency score (FC Schalke 04)
- Curious DMU is the case of Borussia Dortmund, which has just 1 pts less than the best observation in VarPts. However, it is the second-worst in the efficiency score. This rank can be due to both high %NPMin or SqValue
- Bayern München is clearly penalized by having negative VarPts and the highest SqValue

Table 6 - DEA-VRS results of Bundesliga 2019 with %NPMin and SqValue

DMU	%NPMin	SqValue	VarPts	Efficiency
1. FSV Mainz 05	0,4209	86,6	7	1
FC Augsburg	0,154	100,95	-9	1
Hertha BSC	0,1083	125,18	0	1
RB Leipzig	0,1144	314,45	13	1
SV Werder Bremen	0,2472	110,85	11	1
VfL Wolfsburg	0,2354	145,53	22	1
SC Freiburg	0,3036	94,28	0	0,9896
Hannover 96	0,3482	98,28	-18	0,9336
FC Bayern München	0,1319	835,55	-6	0,8215
Borussia Monchengladbach	0,1948	191,53	8	0,7653
Eintracht Frankfurt	0,3654	139,55	5	0,7328
Bayer 04 Leverkusen	0,1585	359,35	3	0,6922
VfB Stuttgart	0,353	159,45	-23	0,6122
TSG 1899 Hoffenheim	0,2429	188,63	-4	0,5892
Borussia Dortmund	0,3799	376,8	21	0,5842
FC Schalke 04	0,3367	243,9	-30	0,4323

Table 7 - DEA-VRS results of Bundesliga 2019 with %NPMin and PFielded

DMU	%NPMin	PFielded	VarPts	Efficiency
Bayer 04 Leverkusen	0,1585	21	3	1
Borussia Dortmund	0,3799	23	21	1
Hertha BSC	0,1083	26	0	1
RB Leipzig	0,1144	23	13	1
VfL Wolfsburg	0,2354	25	22	1
Borussia Monchengladbach	0,1948	22	8	0,9865
FC Bayern München	0,1319	24	-6	0,9402
SV Werder Bremen	0,2472	24	11	0,9187
VfB Stuttgart	0,353	26	-23	0,8077
1, FSV Mainz 05	0,4209	28	7	0,7659
FC Augsburg	0,154	30	-9	0,7621
SC Freiburg	0,3036	28	0	0,75
TSG 1899 Hoffenheim	0,2429	29	-4	0,7241
Eintracht Frankfurt	0,3654	30	5	0,7074
FC Schalke 04	0,3367	32	-30	0,6563
Hannover 96	0,3482	32	-18	0,6563

Next, all the results from every league were put together and every DMU with %NPMin until 20% were aggregated. Until this point, the results were oriented toward efficiency. The purpose is to observe how the variation of points was when the club is considered to have squad stability from one year to another. Consequently, it is also essential to analyse how a club's stability influences the performance compared to the previous season.

Tables 8 and 9 display the sum and average of variation of points in which DMUs have less than 20% of %NPMin.

Table 8 - Sum of Variation of Points with %NPMin below 20%

Sum of VarPts with %NPMin < 20%			
	2019	2018	2017
English Premier League	-32	-46	-13
Bundesliga	9	20	-17
Ligue 1	-30	-	-7
La Liga	-14	-36	-7
Liga NOS	-3	-	-6
Serie A	9	2	-3

Table 9 - Average of Variation of Points with %NPMin below 20%

Average of Varpts with %NPMin < 20%			
	2019	2018	2017
English Premier League	-6,4	-11,5	-3,25
Bundesliga	1,5	20	-5,67
Ligue 1	-15	-	-3,5
La Liga	-4,67	-12	-1,75
Liga NOS	-3	-	-6
Serie A	9	1	-1,5

By analysing tables 8 and 9, it is noticeable that only Bundesliga and Serie A have positive values of points variation. Thus, from these results, it is noticeable that a stable squad is not a direct sign of having better performance than the previous season, as in most of the leagues shown above. Even in terms of efficiency:

- **Premier League:** from 13 observations with less than 20% of %NPMin, 4 (31%) were assigned with maximum efficiency.
- **Bundesliga:** from 10 observations with less than 20% of %NPMin, 4 (40%) were assigned with maximum efficiency.
- **Ligue 1:** from 4 observations with less than 20% of %NPMin, 3 (75%) were assigned with maximum efficiency.
- **La Liga:** from 10 observations with less than 20% of %NPMin, 4 (40%) were assigned with maximum efficiency.
- **Liga NOS:** from 2 observations with less than 20% of %NPMin, 2 (100%) were assigned with maximum efficiency.
- **Serie A:** from 5 observations with less than 20% of %NPMin, 3 (60%) were assigned with maximum efficiency.

From these percentages an average of 58% of the cases with less than 20% of new players minutes have the maximum efficiency score. Consequently, these results confirm what was concluded earlier regarding the relation between %NPMin and VarPts – generally, low values of %NPMin results in higher efficiency scores.

On the other hand, these results also showed that the clubs considered a stable squad (since they did not change more than 20% of the team) had a decline in performance compared to the previous season. However, it is essential to notice that this decline in performance (this case related in terms of points) can be misjudged since a team can have fewer points compared with the previous season even though it can win the league. This is one of the study's limitations since the team accomplished the season goal even with a negative variation of points.

6.2 Linear Regression results

As the name indicates, this chapter aims to address the results of the linear regression model done on the STATA software. The main variables used were the same as before on the DEA model. However, Goals Scored (GScored), as a control variable, and the interaction between %NPMin and PFielded were also included. All the models were performed with the league and the year as control variables, using fixed effects. In other words, even though DMUs were entirely considered, only DMUs from the same year and the same league are put together. This avoids performing models with DMUs that do not play against each other at any point of the season since they are from different leagues. It also avoids including the same team from different seasons on the same model.

Table 10 - Linear Regression Models with one input variable

Input Variables	One variable models	Output variable		
		1. VarPts	2. VarPts	3. VarPts
	%NPMin	10,086 (-1,28)		
	SqValue		0,004 (-0,48)	
	PFielded			**-1,492** (-4,59)
	β_0	-5,175 (-1,29)	-0,567 (-0,27)	**44,514** (-4,42)
	Adjusted R-squared	0,1	0,1	0,16
	F-statistic	3,49	3,43	7,24

*** p<0,1**

**** p<0,05**

***** p<0,01**

Numbers in brackets are the standard deviation

Table 10 displays the models with only the main explanatory variables performed individually and the output variables. Model number 1 was performed with %NPMin as the unique independent variable and VarPts as the output variable. As seen from table 10 and having a close look at the values of the three significance levels, it is notorious that the first model is not significant at any level of the study. Consequently, this model has no explanatory power for the variables in the study. Despite this conclusion, one important topic to observe is that %NPMin positively influences the output variable, VarPts, opposing what was expected and opposing what was supported by DEA models. The second model was essential to establish the squad value as a control value since it has a minor influence on VarPts.

On the other hand, SqValue is critical to be in the model in order to distinguish big and small clubs. Also, this model and the first are not significant and have no explanatory power at any level of significance. Finally, the third model concludes that by increasing one player in the total number of players fielded across the season, VarPts was affected negatively. Thus, the increase in the number of players fielded makes the VarPts decrease gradually. Furthermore, this model is the only one, from these single regression models, that showed significance at any level. There is a 99,99% of confidence, in this case, that the null hypothesis can be rejected, and that

the R-squared is not zero. Consequently, it can be guaranteed that the model has some explanatory power for the values in the study.

Table 11 - Linear Regression Models with two input variables

Input Variables	Two variable models	Output variable	
		4. VarPts	**5. VarPts**
	%NPMin	**13,499** (-1,69)	**20,957** (-2,75)
	SqValue	0,007 (-0,84)	
	PFielded		**-1,739** (-5,41)
	β0	-6,906 (-1,65)	**42,229** (-4,22)
	Adjusted R-squared	0,11	0,18
	F-statistic	3,2	7,66

*** p<0,1**

**** p<0,05**

***** p<0,01**

Numbers in brackets are the standard deviation

Following the analyses of the single regression models, the next step was to perform models with groups of two explanatory variables, however both with %NPMin since it is the focus of study in this masters' dissertation. One important topic is that %NPMin, contrarily to what happened earlier in model 1, showed significance on the lower level ($p < 0,1$). The fourth model was performed together with %NPMin and SqValue. As it was observable in model number 2, SqValue has no power in the model. Even though model 4 shows no tremendous explanatory power to predict the output variable.

On the other hand, model 5 shows excellent results, recognizing that %NPMin and PFielded may have an interaction effect that was not considered. If the model is interpreted with the two variables separately, the conclusion is that one unit of %NPMin affects the VarPts positively. Consequently, it would be positive if a team had a more significant percentage of new player minutes in comparison with minutes of long last players. However, the variable PFielded shows that each player fielded (either new or long last) negatively affect the points. In fact, there is evidence that these two variables have an interaction effect that is not being considered. This interaction effect will be analysed in table 12 below. In addition, another important topic is that model 5 is significant in all variables and parameters at the highest significance level, which concludes a better explanatory power.

Table 12 - Linear regression models with three or more input variables

	Three or more variable models	Output variable		
		6. VarPts	**7. VarPts**	**8. VarPts**
Input Variables	**%NPMin**	**23,534** (-3,24)	**27,14** (-3,98)	**208,395** (-3,91)
	SqValue	0,005 (-0,63)	-0,012 (-1,03)	-0,011 (-0,93)
	PFielded	**-1,721** (-5,22)	**-1,463** (-4,35)	0,807 (-1,08)
	GScored		**0,334** (-3,75)	**0,332** (-3,88)
	%NPMin . PFielded			**-6,488** (-3,5)
	β0	**40,371** (-3,76)	16,116 (-1,26)	**-45,868** (-2,07)
	Adjusted R-squared	0,18	0,25	0,28
	F-statistic	7,15	9,12	9,08

*** p<0,1**

**** p<0,05**

***** p<0,01**

Numbers in brackets are the standard deviation

The next step after performing models 4 and 5 was to include more variables into the study. Model 6 was performed with the three variables simultaneously – NPMin, SqValue and PFielded. The results were robust and continued with the same line of thought as the previous ones: %NPMin affecting the variation of points positively; SqValue as a control variable; and PFielded negatively affecting the output variable. Note that either %NPMin, PFielded and β_0 have significant power on the lowest p-value – resulting in a robust explanatory model.

$$\mathbf{VarPts} = 23{,}53\mathbf{NPMin} + 0{,}005\mathbf{SqValue} - 1{,}72\mathbf{PFielded} + 40{,}37 \qquad (14)$$

In equation 14, it is possible to observe the model in a more straightforward way in order to interpret the parameters of each variable. In this model, %NPMin affects the variation of points positively, and the number of players fielded the opposite, which was first noticed in model 5. One percentual point increase on %NPMin influences a increase of 23,53 points compared with last season result. Contrarily, for each new PFielded (without considering if he was already on the team or not), the number of points suffers a decrease of 1,72 compared to the previous season. Moreover, these results were concluding the opposite of the previous results taken from DEA models. Naturally, taking a deeper look at both these results may not be coherent since they have an interaction value that may not be considered. Consequently, it became more important to perform a model considering the possible interaction effect between %NPMin and PFielded – model 8.

Before performing model 8, and with the justification of increasing the final model robustness, an additional control variable was included – GScored – which helped to understand whether the model was well built, since GScored also affects VarPts positively as expected. The result of this variable was easily predictable since it is common sense that when a team scores more goals it has a higher probability of having a higher point score than the previous season.

$$\mathbf{VarPts} = 208{,}4\mathbf{NPMin} - 0{,}011\mathbf{SqValue} + 0{,}81\mathbf{PFielded} + 0{,}33\mathbf{GScored} - 6{,}49\mathbf{NPMin.PFielded} - 45{,}87 \quad (15)$$

In this last model, the primary variable of interest was included, as mentioned before. The conclusions from all the models performed, until model 8, pointed out that %NPMin and PFielded would have an interaction. Consequently, this model has an additional variable related to that interaction effect – NPMin.Pfielded. Therefore, on this model 8, there are three variables of interest together with two control variables, as explained earlier. Equations 16 and 17 help to better interpret the interaction effect results between the two main variables, representing the impact of %NPMin on VarPts and the impact of PFielded on the same variable, respectively.

$$\frac{\Delta \mathrm{Pts}}{\%\mathbf{NPMin}} = 208{,}4 - 6{,}49\mathbf{PFielded} \quad (16)$$

From equation 16, it is possible to understand that a minor variation of points happens when there is an increase in %NPMin and a higher PFielded. This was concluded earlier and is now proved with this model. Furthermore, the equation has a robust explanatory power since both parameters involved are significant on the highest interval possible. The same does not happen in equation 17.

$$\frac{\Delta \mathrm{Pts}}{\mathbf{PFielded}} = 0{,}81 - 6{,}49\mathbf{NPMin} \quad (17)$$

With a similar purpose to equation 16, the objective this time was to understand the impact of PFielded, including the interaction effect on the output variable. However, the parameter of PFielded on the primary equation of the model has no significant power, which makes this equation have less explanatory power than equation 16. In this case, an increase in %NPMin would affect, on a large scale, the variation of points negatively.

6.3 Chapter conclusions

Initially, expectations were aligned with what had been published in the literature, concluding that when a team changes a significant number of players from one season to the next, performance is harmed. Subsequently, from the results of DEA models, no direct conclusions were possible to draw. It was only possible to see an evident tendency line reflecting the relation between %NPMin and VarPts. Finally, linear regression models, more precisely model number 8, allowed to prove that stability affects the team's performance. Clubs that have fewer new players fielded are more likely to have a better performance than before.

7. Conclusions and Future Considerations

The world of modern football is related a lot to the enterprise environment. In the last decades, football clubs became and continue to evolve into huge companies to manage and need to optimise their more valuable assets – the players.

Football suffered many changes in comparison with some decades ago. Today, the technology is a lot incorporated in studies regarding teams, players, performance in general, and many statistical methods used in the football world. In this dissertation, the impact of changes in the roster on team performance is studied.

This chapter aims to show the main conclusions taken from this dissertation, both from the models used and the studies analysed in the literature review. In chapter 3, it was showed an overview of various studies regarding many subjects: Productivity, Effectiveness, Efficiency, Job turnover, Manager turnover, Squad stability, Player turnover and Team performance. After that analysis of the literature, both models used were explained: DEA and Multiple Linear Regression.

In this experimental stage of the dissertation, the first step was to decide and collect the data needed for the study. The decision came in the direction of the six most important leagues: EPL, La Liga, Bundesliga, Serie A, Ligue 1 and Liga NOS. Despite knowing the different competitive levels between the leagues since the models were performed by league and season, comparing teams of different realities was more avoided. Consequently, the results were analysed individually for each league and season.

Regarding DEA models, the main difficulty was to have strict conclusions with the results. However, it was possible to analyse a less deep look and see a tendency line in all leagues every season. All observations showed the same tendency line, only differing on the line slope. On the %NPMin vs Efficiency analysis, it was clear that a higher %NPMin relates to a lower efficiency rating. Bundesliga was the league where this relationship was more notable over the three years of analysis, only surpassed by the Premier League in 2017. In other words, all seasons proved that when the team's stability decreases, the value of efficiency also decreases. From DEA models, another strong tendency was analysed even though not so important as the primary variable of the study (%NPMin), which was the behaviour of PFielded vs Efficiency. The variable PFielded had similar results as the primary variable in that when a team has a higher number of PFielded (less stability), the efficiency rating is low.

The next step was to perform the linear regression model with different variables and some new ones to have a more robust model. The first expectations were that the results would prove and support the previous DEA ones. Both models supported the same conclusions; however, these regression models were critical to analyse and understand the interaction effect between %NPMin and PFielded. This interaction effect was proved to hurt the overall team performance, which supported the results obtained before by the DEA models. Consequently, the main goal of this dissertation was accomplished and was aligned with the existent literature.

7.1 Limitations and Difficulties

Throughout the study, some topics limited the range of the study and, consequently, can be considered as an imperfection of this dissertation:

1. Even though season 2019/2020 was already finished, it was decided not to be included since all leagues were interrupted due to the pandemic of covid-19. All leagues except for Ligue 1 have resumed, although in a very different environment which could influence all results.
2. Promoted and relegated were not considered since it was not possible to collect data from two consecutive seasons.
3. DEA models showed in most of the results a high number of maximum efficiency ratings.
4. In the models there are not an evident conclusion that is the turnover influencing performance or the opposite. On other words, it is not evident if it is the squad stability changing the performance from season to season or a bad performance influencing the changing on the squads.

In terms of difficulties across the process of concluding the dissertation, important to highlight two of them:

1. A predefined dataset was not ready. Consequently, it was made by organic search in TransferMarkt, SoccerStats and ZeroZero, which was long time consuming process.
2. No previous knowledge of the models applied on the dissertation.

7.2 Future Considerations

Every study and dissertation can help giving insights for further investigations regarding the same subject or investigations that will use the same models applied here. For further investigations, here are the recommendations:

1. New variables
2. More leagues
3. Expand to other sports
4. Presidents and Directors turnover instead of player turnover
5. Apply optimization tools and models

References

Allen, M. P., Panian, S. K., & Lotz, R. E. (1979). *Managerial Succession and Organizational Performance : A Recalcitrant Problem Revisited*. *24*(2), 167–180. http://www.jstor.org/stable/2392492

Álvarez, I. C., Barbero, J., Zofío, J. L., & Barbero, J. (2016). *A Data Envelopment Analysis Toolbox for MATLAB*.

Audas, R., Dobson, S., & Goddard, J. (1999). Organizational performance and managerial turnover. *Managerial and Decision Economics*, *20*(6), 305–318. https://doi.org/10.1002/(SICI)1099-1468(199909)20:6<305::AID-MDE933>3.0.CO;2-O

Audas, R., Dobson, S., & Goddard, J. (2002). The impact of managerial change on team performance in professional sports. *Journal of Economics and Business*, *54*(6), 633–650. https://doi.org/10.1016/S0148-6195(02)00120-0

Bachan, R., Reilly, B., & Witt, R. (2008). The hazard of being an English football league manager: Empirical estimates for three recent league seasons. *Journal of the Operational Research Society*, *59*(7), 884–891. https://doi.org/10.1057/palgrave.jors.2602408

Banker, R., Charnes, A., & Cooper, W. (1984). Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Management Science*, *30*(9), 1078–1092.

Barnard, M., Boor, S., Winn, C., Wood, C., & Wray, I. (2019). Annual Review of Football Finance 2019. *Deloitte Annual Review of Football Finance 2019*, *May*, 40.

Barros, C., & Athanassiou, M. (2004). Efficiency in European seaports with DEA: evidence from Greece and Portugal. *Maritime Economics & Logistics*, *6*(2), 122–140.

Barros, C. P., Frick, B., & Passos, J. (2009). Coaching for survival: The hazards of head coach careers in the German “Bundesliga.” *Applied Economics*, *41*(25), 3303–3311. https://doi.org/10.1080/00036840701721455

Batt, R. (2002). Managing customer services: Human resource practices, quit rates, and sales growth. *Academy of Management Journal*, *45*(3), 587–597. https://doi.org/10.2307/3069383

Bernard, Y., Godard, L., & Zouaoui, M. (2018). The Effect of CEOs’ Turnover on the Corporate Sustainability Performance of French Firms. *Journal of Business Ethics*, *150*(4), 1049–1069. https://doi.org/10.1007/s10551-016-3178-7

Boscá, J. E., Liern, V., Martínez, A., & Sala, R. (2009). Increasing offensive or defensive efficiency? An analysis of Italian and Spanish football. *Omega*, *37*(1), 63–78. https://doi.org/10.1016/j.omega.2006.08.002

Bremer, M. (2012). *Multiple Linear Regression*. Mezeylab. http://mezeylab.cb.bscb.cornell.edu/labmembers/documents/supplement 5 - multiple regression.pdf

Brown, M. C. (1982). *Administrative Succession and Organizational Performance : The Succession Effect*. *27*(1), 1–16. http://www.jstor.org/stable/2392543

Carmichael, F., Rossi, G., & Thomas, D. (2014). Production, Efficiency, and Corruption in Italian Serie A Football. *Journal of Sports Economics*, *18*(1), 34–57. https://doi.org/10.1177/1527002514551802

Carvalho, M., & Marques, R. (2007). *Performance Evaluation of the Portuguese Seaports Evaluation in the*

European Context. INSTITUTO SUPERIOR TÉCNICO - Universidade Técnica de Lisboa.

Charnes, A., Cooper, W. W., & Rhodes, E. (1978). Measuring the efficiency of decision-making units. *European Journal of Operational Research*, *2*(6), 429–444. https://doi.org/10.1016/0377-2217(78)90138-8

Charnes, Abraham, Cooper, W. W., Lewin, A. Y., & Seiford, L. M. (1984). *Data Envelopment Analysis: Theory, Methodology, and Applications* (1st ed.). Springer Netherlands. https://doi.org/10.1007/978-94-011-0637-5

Chen, L. F., Hsiao, C. H., & Tsai, C. F. (2010). Three-stage-DEA model selections and managerial decision. *African Journal of Business Management*, *4*(14), 3046–3055.

CIES. (2016a). *Monthly report 19 - Demographic study of football in Europe*. CIES Football Observatory. https://football-observatory.com/IMG/sites/mr/mr19/en/

CIES. (2016b). *Weekly post 165 - Squad stability: the european league rankings*. CIES Football Observatory. https://football-observatory.com/IMG/sites/b5wp/2016/165/en/

CIES. (2018). *Monthly report 34 - The importance of squad stability: evidence from European football*. CIES Football Observatory. https://football-observatory.com/IMG/pdf/mr34en.pdf

CIES. (2019). *Monthly Report 49 - The demographics of football in the European labour market*. CIES Football Observatory. https://football-observatory.com/IMG/sites/mr/mr49/en/

Coelli, T. (1996). *A Guide to DEAP Version 2.1: A Data Envelopment Analysis (Computer) Program*. http://www.une.edu.au/econometrics/cepa.htm CEPA

Cohen, S. G., & Bailey, D. E. (1997). What makes teams work: Group effectiveness research from the shop floor to the executive suite. In *Journal of Management* (Vol. 23, Issue 3). https://doi.org/10.1177/014920639702300303

Cook, W., & Zhu, J. (2005). Building Performance Standards into Data Envelopment Analysis Structures. *IIE Transactions*, *37*(3), 267–275.

Cooper, W., Seiford, L., & Zhu, J. (2006). International Series in Operations Research & Management Science. In *Handbook on Data Envelopment Analysis* (Second Edi, p. 498). Springer US. https://doi.org/10.1007/978-1-4419-6151-8

Cullinane, K., Song, D.-W., Ji, P., & Wang, T.-F. (2004). An Application of DEA Windows Analysis to Container Port Production Efficiency. *Review of Network Economics*, *3*(2), 184–206. https://doi.org/10.2202/1446-9022.1050

d'Addona, S., & Kind, A. (2014). Forced Manager Turnovers in English Soccer Leagues: A Long-Term Perspective. *Journal of Sports Economics*, *15*(2), 150–179. https://doi.org/10.1177/1527002512447803

Denis, D. J., & Denis, D. K. (1995). Performance Changes Following Top Management Dismissals. *The Journal of Finance*, *50*(4), 1029–1057. https://doi.org/10.1111/j.1540-6261.1995.tb04049.x

Denis, D. J., Denis, D. K., & Sarin, A. (1997). Agency Problems, Equity Ownership, and Corporate Diversification. *CFA Digest*, *27*(4), 49–51. https://doi.org/10.2469/dig.v27.n4.171

Dobson, S., & Goddard, J. (2001). The Economics of Football. In *The Economics of Football* (Issue January 2001). https://doi.org/10.1017/cbo9780511493225

Drake, L. M., & Simper, R. (2005). Police efficiency in offences cleared: An analysis of english “Basic Command Units.” *International Review of Law and Economics*, *25*(2), 186–208. https://doi.org/10.1016/j.irle.2005.06.003

Eitzen, D. S., & Yetman, N. R. (1972). Managerial Change, Longevity, and Organizational Effectiveness. *Administrative Science Quarterly*, *17*(1), 110. https://doi.org/10.2307/2392099

Eriksson, T., & Ortega, J. (2006). The adoption of job rotation: Testing the theories. *Industrial and Labor Relations Review*, *59*(4), 653–666. https://doi.org/10.1177/001979390605900407

Farrell, K. A., & Whidbee, D. A. (2002). Monitoring by the financial press and forced CEO turnover. *Journal of Banking and Finance*, *26*(12), 2249–2276. https://doi.org/10.1016/S0378-4266(01)00183-2

Farrell, K. A., & Whidbee, D. A. (2003). Impact of firm performance expectations on CEO turnover and replacement decisions. *Journal of Accounting and Economics*, *36*(1-3 SPEC. ISS.), 165–196. https://doi.org/10.1016/j.jacceco.2003.09.001

Farrell, M. J. (1957). The Measurement of Productive Efficiency. *Journal of the Royal Statistical Society. Series A (General)*, *120*(3), 253–290. http://goo.gl/AFhm2N

Fee, C. E., & Hadlock, C. J. (2004). Management turnover across the corporate hierarchy. *Journal of Accounting and Economics*, *37*(1), 3–38. https://doi.org/10.1016/j.jacceco.2003.11.003

FIFA. (2006). *Commentary on the Regulations for the Status and Transfer of Players*. 76. http://resources.fifa.com/mm/document/affederation/administration/51/56/07/transfer_commentary_06_en_1843.pdf

Fizel, J. L., & D’Itri, M. P. (1997). Managerial efficiency, managerial succession and organizational performance. *Managerial and Decision Economics*, *18*(4), 295–308. https://doi.org/10.1002/(SICI)1099-1468(199706)18:4<295::AID-MDE828>3.0.CO;2-W

Football-Stadiums.co.uk. (2020). *Football Player Transfers Explained*. https://www.football-stadiums.co.uk/articles/football-player-transfers-explained/

Football World Rankings. (2018). *FOOTBALL COACH TURNOVER INDEX 2018*. www.clubworldranking.com

Forbes. (2019a). *The Business Of Soccer - 2019 Most Valuable Clubs Ranking*. https://www.forbes.com/soccer-valuations/list/#tab:overall

Forbes. (2019b). *The World’s Most Valuable Soccer Teams 2019*. https://www.forbes.com/sites/mikeozanian/2019/05/29/the-worlds-most-valuable-soccer-teams-2019/#1c61300540d6

Forbes. (2020). *The World’s Highest-Paid Athletes List - Forbes*. Forbes.Com. http://www.forbes.com/athletes/list/

Frick, B., Barros, C. P., & Prinz, J. (2010). Analysing head coach dismissals in the German “Bundesliga” with a mixed logit approach. *European Journal of Operational Research*, *200*(1), 151–159. https://doi.org/10.1016/j.ejor.2008.11.048

Frost, J. (2018). *How to Interpret P-values and Coefficients in Regression Analysis*. Statistics by Jim. https://statisticsbyjim.com/regression/interpret-coefficients-p-values-regression/

Gammage, K. L., Carron, A. V., & Estabrooks, P. A. (2001). Team Cohesion and Individual Productivity:

The Influence of the Norm for Productivity and the Identifiability. *Small Group Research*, *32*(1), 3–18.

Gamson, W. a, & Scotch, N. a. (1963). Scapegoating in Baseball AND DEBATES Scapegoating in Baseball. *American Journal of Sociology*, *70*(1), 69–72.

Grusky, O. (1963). Managerial Succession and Organizational Effectiveness. *American Journal of Sociology*, *69*(1), 21–31. https://doi.org/10.1086/223507

Gully, S. M., Devine, D. J., & Whitney, D. J. (1995). *A Meta-Analysis of Cohesion and Performance: Effects of Level of Analysis and Task Interdependence*. *26*(4), 497–520.

Guthrie, J. P. (2001). High-Involvement Work Practices, Turnover, and Productivity: Evidence From New Zealand. *Academy of Management Journal*, *44*(1), 180–190. http://www.jstor.org/stable/3069345

Haas, D. J. (2003). Productive Efficiency of English Football Teams - A Data Envelopment Analysis Approach. *Managerial and Decision Economics*, *24*, 403–410. https://doi.org/10.1002/mde.1105

Haas, D. J., Kocher, M. G., & Sutter, M. (2004). Measuring Efficiency of German Football Teams by Data Envelopment Analysis. *Central European Journal of Operations Research*, *12*(3), 251–268. http://search.proquest.com/docview/195543721?pq-origsite=gscholar

Hancock, J. I., Allen, D. G., Bosco, F. A., McDaniel, K. R., & Pierce, C. A. (2013). Meta-Analytic Review of Employee Turnover as a Predictor of Firm Performance. *Journal of Management*, *39*(3), 573–603. https://doi.org/10.1177/0149206311424943

Hancock, J. I., Allen, D. G., & Soelberg, C. (2017). Collective turnover: An expanded meta-analytic exploration and comparison. *Human Resource Management Review*, *27*(1), 61–86. https://doi.org/10.1016/j.hrmr.2016.06.003

Hausknecht, J. P., & Trevor, C. O. (2011). Collective turnover at the group, unit and organizational levels: Evidence, issues, and implications. *Journal of Management*, *37*(1), 352–388.

Heavey, A. L., Holwerda, J. A., & Hausknecht, J. P. (2013). Causes and consequences of collective turnover: A meta-analytic review. *Journal of Applied Psychology*, *98*(3), 412–453. https://doi.org/10.1037/a0032380

Hinkin, T. R., & Tracey, J. B. (2000). The Cost of Turnover. *Cornell Hotel and Restaurant Administration Quarterly*, *41*(3), 14–21. https://doi.org/10.1177/001088040004100313

Hu, J. L., Chiu, C.-N., Shieh, H.-S., & Huang, C.-H. (2010). A stochastic cost efficiency analysis of international tourist hotels in Taiwan. *International Journal of Hospitality Management*, *29*(1), 99–107.

Huselid, M. A. (1995). The Impact Of Human Resource Management Practices On Turnover, Productivity, And Corporate Financial Performance. *Human Resource Management*, *3*(3).

Huson, M. R., Parrino, R., & Starks, L. T. (2001). Internal monitoring mechanisms and CEO turnover: A long-term perspective. *Journal of Finance*, *56*(6), 2265–2297. https://doi.org/10.1111/0022-1082.00405

Kampkötter, P., Harbring, C., & Sliwka, D. (2018). Job rotation and employee performance–evidence from a longitudinal study in the financial services industry. *International Journal of Human Resource Management*, *29*(10), 1709–1735. https://doi.org/10.1080/09585192.2016.1209227

Kim, M., & Harris, T. (2008). Efficiency analysis of the US biotechnology industry: clustering enhances

productivity. *The Journal of Agrobiotechnology Management & Economics*, *12*(3, 4).

Koning, R. H. (2003). An econometric evaluation of the effect of firing a coach on team performance. *Applied Economics*, *35*(5), 555–564. https://doi.org/10.1080/0003684022000015946

Kounetas, K. (2014). Greek foot ball clubs' efficiency before and after Euro 2004 Victory: a bootstrap approach. *Central European Journal of Operations Research*, *22*(4), 623–645. https://doi.org/10.1007/s10100-013-0288-5

Koys, D. J. (2001). The effects of employee satisfaction, organizational citizenship behavior, and turnover on organizational effectiveness: A unit-level, longitudinal study. *Personnel Psychology*, *54*(1), 101–114. https://doi.org/10.1111/j.1744-6570.2001.tb00087.x

L. Davis, J., Fodor, A., E. Pfahl, M., & Stoner, J. (2014). Team interdependence and turnover: evidence from the NFL. *American Journal of Business*, *29*(3/4), 276–292. https://doi.org/10.1108/ajb-02-2014-0009

Lieberson, S., & Connor, J. F. O. (1972). *Leadership and organizational performance: A study of large corporations*. *37*(2), 117–130.

Lukomski, J. (2020). On the finalisation of international football transfers and professional football players ' contracts. *The International Sports Law Journal*, *0123456789*. https://doi.org/10.1007/s40318-020-00170-y

Madina Tatraeva. (2020). *Buy-out clauses in football contracts*. ILaw.Co.Uk. https://www.ilaw.co.uk/news/buy-out-clauses-in-football-contracts

Marill, K. (2004). Advanced Statistics: Linear Regression, Part II: Multiple Linear Regression. *Academic Emergency Medicine : Official Journal of the Society for Academic Emergency Medicine*, *11*, 94–102. https://doi.org/10.1197/j.aem.2003.09.005

Marques, R. C. (2011). Advanced Operations Research Performance Evaluation. In *Slides das Aulas de Complementos de Investigação Operacional, Instituto Superior Técnico*.

Marques, R., & Silva, D. (2006). Inferência estatística dos estimadores de eficiência obtidos com a técnica fronteira não paramétrica de DEA: Uma metodologia de Bootstrap. *Associação Portuguesa de Investigação Operacional*, *26*(1), 89–110.

Mbuvi, D., De Witte, K., & Perelman, S. (2012). Urban water sector performance in Africa: A step-wise bias-corrected efficiency and effectiveness analysis. *Utilities Policy*, *22*, 31–40. https://doi.org/10.1016/j.jup.2012.02.004

Michele Kacmar, K., Andrews, M. C., Van Rooy, D. L., Chris Steilberg, R., & Cerrone, S. (2006). Sure everyone can be replaced . . . But at what cost? Turnover as a predictor of unit-level performance. *Academy of Management Journal*, *49*(1), 133–144. https://doi.org/10.5465/AMJ.2006.20785670

Mohr, D. C., Young, G. J., & Burgess, J. F. (2012). Employee turnover and operational performance: The moderating effect of group-oriented organisational culture. *Human Resource Management Journal*, *22*(2), 216–233. https://doi.org/10.1111/j.1748-8583.2010.00159.x

Mohsin, A., Lengler, J., & Aguzzoli, R. (2015). Staff turnover in hotels: Exploring the quadratic and linear relationships. *Tourism Management*, *51*, 35–48. https://doi.org/10.1016/j.tourman.2015.04.005

Morse, A. L., Shapiro, S. L., Mcevoy, C. D., & Rascher, D. A. (2008). The effects of roster turnover on

demand in the National Basketball Association. *International Journal of Sport Finance*, *3*(1), 8–18. https://doi.org/10.2139/ssrn.1690885

Mortimer, D., & Peacock, S. (2002). Hospital Efficiency Measurement : Simple Ratios vs Frontier Methods. *Centre for Health Program Evaluation*, *135*.

Mouzas, S. (2006). Efficiency versus effectiveness in business networks. *Journal of Business Research*, *59*(10–11), 1124–1132. https://doi.org/10.1016/j.jbusres.2006.09.018

Muehlheusser, G., Schneemann, S., & Sliwka, D. (2016). The impact of managerial change on performance: The role of team heterogeneity. *Economic Inquiry*, *54*(2), 1128–1149. https://doi.org/10.1111/ecin.12285

O'Fallon, M. J., & Rutherford, D. G. (2010). *Hotel Management and Operations* (5th ed.). John Wiley & Sons, Inc. https://books.google.pt/books?hl=pt-PT&lr=&id=G0AnmDecKpQC&oi=fnd&pg=PR15&ots=sp3EqF9GKA&sig=0wz0AoK7KmvlD8JPkR2iBgdDeJ4&redir_esc=y#v=onepage&q&f=false

Ogbonna, E., & Harris, L. C. (2002). Managing organisational culture: Insights from the hospitality industry. *Human Resource Management Journal*, *12*(1), 33–53. https://doi.org/10.1111/j.1748-8583.2002.tb00056.x

Park, T. Y., & Shaw, J. D. (2013). Turnover rates and organizational performance: A meta-analysis. *Journal of Applied Psychology*, *98*(2), 268–309. https://doi.org/10.1037/a0030723

Parrino, R. (1997). JOURNAL OF Financial ECONOMICS CEO turnover and outside succession A cross-sectional analysis. *Journal of Financial Economics*, *46*, 165–197.

Poulsen, R. (2000). Should He Stay or Should He Go? Estimating the Effect of Firing the Manager in Soccer. In *Chance* (Vol. 13, Issue 2, pp. 29–32). https://doi.org/10.1080/09332480.2000.10542204

Reinganum, M. R. (1985). The Effect of Executive Succession on Stockholders Wealth of Large Corporations. *Academy of Management Proceedings*, *1985*(1), 36–40. https://doi.org/10.5465/ambpp.1986.4978498

Renaud, O., & Victoria-Feser, M.-P. (2010). A robust coefficient of determination for regression. *Journal of Statistical Planning and Inference*, *140*(7), 1852–1862. https://doi.org/10.1016/j.jspi.2010.01.008

Salomo, S., & Teichmann, K. (2000). The Relationship of Performance and Managerial Succession in the German Premier Football League. *European Journal of Sport Management*, *7*, 99–119.

Sarafidis, V. (2002). An Assessment of Comparative Efficiency Measurement Techniques. *Europe Economics*, 1–21.

Schneider, A., Hommel, G., & Blettner, M. (2010). Linear Regression Analysis. *Deutsches Arzteblatt International*, *107*(44), 776–782. https://doi.org/10.3238/arztebl.2010.0776

Scully, G. W. (1994). Managerial efficiency and survivability in professional team sports. *Managerial and Decision Economics*, *15*(5), 403–411. https://doi.org/10.1002/mde.4090150503

Shook, A. (2016). *The Determinants of Forced Manager Turnovers in Major League Soccer*.

Siebert, W. S., & Zubanov, N. (2009). Searching for The Optimal Level of Employee Turnover: A Study of a Large U.K. Retail Organization. *Academy of Management Journal*, *52*(2).

Siegel, A. F. (2016). Practical Business Statistics. In A. F. Siegel (Ed.), *Practical Business Statistics* (Seventh Ed, pp. 299–354). Academic Press. https://doi.org/10.1016/B978-0-12-804250-2.00011-0

Simar, L. (1992). Estimating Efficiencies from Frontier Models with Panel Data: A Comparison of Parametric, Non-Parametric and Semi-Parametric Methods with Bootstrapping. *Journal of Productivity Analysis*, *3*(1/2), 171–203. http://www.jstor.org/stable/41770579

SoccerStats. (2020). *Clubs Statistics*. https://www.soccerstats.com/

Sorensen, K. L., Ng, T. W. H., Eby, L. T., & Feldman, D. C. (2007). Determinants of job mobility: A theoretical integration and extension. *Journal of Occupational and Organizational Psychology*, *80*(3), 363–386. https://doi.org/10.1348/096317906X130582

SportsShow. (2020). *Top 10 Most Popular Sports in the World*. https://sportsshow.net/top-10-most-popular-sports-in-the-world/

Statista. (2019). *Soccer players highest transfer fees all-time 2019*. https://www.statista.com/statistics/263304/transfer-fees-the-10-most-expensive-transfers-in-soccer-ever/

Statistics Solution. (2020). *Assumptions of Multiple Linear Regression*. Statistics Solution. https://www.statisticssolutions.com/assumptions-of-multiple-linear-regression/

Stephan Uersfeld. (2017). *Bundesliga study shows players stay at clubs for less than 2 years on average*. Espn.Co.Uk. https://www.espn.co.uk/football/german-bundesliga/story/3208033/bundesliga-study-shows-players-stay-at-clubs-for-less-than-2-years-on-average

talkSPORT. (2015). *Premier League clubs ranked by the average stay of their players*. https://talksport.com/football/343081/premier-league-clubs-ranked-average-stay-their-players-shortest-longest-151112173905/

Tena, J. de D., & Forrest, D. (2007). Within-season dismissal of football coaches: Statistical analysis of causes and consequences. *European Journal of Operational Research*, *181*(1), 362–373. https://doi.org/10.1016/j.ejor.2006.05.024

ter Weel, B. (2011). Does Manager Turnover Improve Firm Performance? Evidence from Dutch Soccer, 1986-2004. *Economist*, *159*(3), 279–303. https://doi.org/10.1007/s10645-010-9157-y

Thanassoulis, E. (2001). *Introduction to the Theory and Application of Data Envelopment Analysis* (1st ed.). Springer US. https://doi.org/10.1007/978-1-4615-1407-7

TotalSportek. (2017). *25 World's Most Popular Sports (Ranked by 13 factors)*. https://www.totalsportek.com/most-popular-sports/

TransferMarkt. (2020a). *Clubs Statistics*. www.transfermarkt.com

TransferMarkt. (2020b). *Most valuable clubs (Detailed view)*. https://www.transfermarkt.com/vereins-statistik/wertvollstemannschaften/marktwertetop?kontinent_id=6&plus=1

UEFA. (2020). *Country coefficients | UEFA Coefficients*. https://www.uefa.com/memberassociations/uefarankings/country/#/yr/2020

van Ours, J. C., & van Tuijl, M. A. (2016). In-season head-coach dismissals and the performance of professional football teams. *Economic Inquiry*, *54*(1), 591–604. https://doi.org/10.1111/ecin.12280

Vincent, L. A. (1959). Définition et mesure de la productivité Journal. *JOURNAL DE LA SOCIÉTÉ STATISTIQUE DE PARIS*, *100*, 35–46.

Weiner, N., & Mahoney, T. A. (1981). A Model of Corporate Performance as a Function of Environmental, Organizational, and Leadership Influences. *Academy of Management Journal*, *24*(3), 453–470. https://doi.org/10.5465/255568

Williams, D., Harris, C., & Parker, J. (2008). I love You - Goodbye: Exit Interviews and Turnover in the New Zealand Hotel Industry. *New Zealand Journal of Employment Relations*, *33*(3), 70–90.

Wooldridge, J. M. (2012). *Introductory econometrics : a modern approach* (5th ed.). South-Western, Cengage Learning.

Wu, T. H., Chen, M. S., & Yeh, J. Y. (2010). Measuring the performance of police forces in Taiwan using data envelopment analysis. *Evaluation and Program Planning*, *33*(3), 246–254.

ZeroZero. (2020). *Minutes played by all Players fielded*. https://www.zerozero.pt/home.php

Appendix A – Correlation Table

	Games played	*Points*	*Victories*	*Draws*	*Losses*	*Goals Scored*	*Goals conceded*	*Goal Average*	*% Clean Sheets*	*% Failed to score*	*Players Fielded*	*Squad value (M €)*	*Sum New players minutes*	*Sum Last year players minutes*	*% New players minutes*	*% Last year players minutes*	*Δ Points*	*Δ Goals Scored*	*Δ Goals conceded*	*Δ % Clean Sheets*	*Δ % Failed to score*	*Δ Players Fielded*	*Δ Squad Value (M €)*	*Δ Sum of New players minutes*	*Δ Sum of Last year players minutes*	*Δ % New players minutes*	*Δ % Last year players minutes*
Games played	1,00																										
Points	0,03	1,00																									
Victories	0,04	0,94	1,00																								
Draws	0,31	-0,47	-0,48	1,00																							
Losses	0,19	-0,68	-0,53	0,16	1,00																						
Goals Scored	0,09	0,73	0,83	-0,16	-0,36	1,00																					
Goals conceded	-0,10	-0,01	-0,25	-0,39	-0,18	-0,41	1,00																				
Goal Average	0,12	0,48	0,66	-0,11	0,14	0,75	-0,66	1,00																			
% Clean Sheets	0,09	0,64	0,70	-0,28	-0,17	0,50	-0,40	0,71	1,00																		
% Failed to score	0,04	-0,70	-0,72	0,35	0,22	-0,74	0,09	-0,67	-0,51	1,00																	
Players Fielded	-0,12	-0,33	-0,32	0,14	0,09	-0,21	-0,02	-0,26	-0,27	0,34	1,00																
Squad value (M €)	0,27	0,56	0,62	-0,29	-0,09	0,56	-0,16	0,57	0,52	-0,52	-0,22	1,00															
Sum New players minutes	0,02	-0,25	-0,28	0,28	0,03	-0,23	-0,09	-0,20	-0,17	0,28	0,36	-0,40	1,00														
Sum Last year players minutes	0,32	0,24	0,27	-0,16	0,04	0,25	0,05	0,24	0,19	-0,26	-0,37	0,46	-0,94	1,00													
% New players minutes	-0,11	-0,25	-0,28	0,23	0,01	-0,25	-0,07	-0,22	-0,18	0,28	0,37	-0,42	0,99	-0,97	1,00												
% Last year players minutes	0,11	0,25	0,28	-0,23	-0,01	0,25	0,08	0,22	0,18	-0,28	-0,38	0,44	-0,99	0,97	-1,00	1,00											
Δ Points	-0,02	0,53	0,37	-0,14	-0,66	0,20	0,22	-0,11	0,14	-0,19	-0,25	0,06	0,03	-0,03	0,03	-0,03	1,00										
Δ Goals Scored	-0,12	0,22	0,22	0,16	-0,41	0,41	-0,21	0,07	-0,02	-0,15	-0,03	-0,05	0,14	-0,17	0,15	-0,14	0,49	1,00									
Δ Goals conceded	-0,09	0,22	0,03	-0,40	-0,26	-0,15	0,73	-0,33	-0,12	-0,14	-0,09	0,01	-0,16	0,12	-0,14	0,15	0,39	-0,22	1,00								
Δ % Clean Sheets	0,07	0,10	0,12	-0,01	0,05	0,03	-0,20	0,23	0,49	-0,02	-0,27	-0,01	0,03	0,00	0,02	-0,03	0,13	-0,04	-0,35	1,00							
Δ % Failed to score	0,07	-0,21	-0,17	0,12	0,13	-0,19	-0,13	-0,10	-0,05	0,50	0,24	0,04	-0,03	0,04	-0,04	0,04	-0,33	-0,44	-0,21	0,04	1,00						
Δ Players Fielded	-0,01	-0,22	-0,19	0,11	0,12	-0,06	-0,02	-0,10	-0,16	0,14	0,67	0,08	0,06	-0,05	0,06	-0,05	-0,31	-0,09	-0,04	-0,26	0,20	1,00					
Δ Squad Value (M €)	0,24	0,46	0,49	-0,16	-0,24	0,43	-0,13	0,33	0,37	-0,34	-0,16	0,79	-0,32	0,39	-0,34	0,36	0,14	0,07	0,01	-0,06	0,03	0,16	1,00				
Δ Sum of New players minutes	0,04	-0,16	-0,18	0,10	0,06	-0,11	0,05	-0,14	-0,12	0,15	0,21	-0,13	0,48	-0,44	0,47	-0,46	0,00	0,08	-0,04	0,01	-0,02	0,12	-0,11	1,00			
Δ Sum of Last year players minutes	-0,02	-0,18	-0,19	-0,03	0,06	-0,20	0,14	-0,21	-0,21	0,18	-0,11	-0,18	-0,26	0,24	-0,25	0,24	-0,01	-0,04	0,13	-0,06	-0,04	-0,12	0,00	0,09	1,00		
Δ % New players minutes	0,03	-0,15	-0,18	0,09	0,06	-0,11	0,05	-0,14	-0,12	0,15	0,21	-0,13	0,48	-0,44	0,47	-0,46	0,00	0,07	-0,04	0,00	-0,01	0,12	-0,12	1,00	0,09	1,00	
Δ % Last year players minutes	-0,03	-0,18	-0,19	-0,04	0,06	-0,20	0,15	-0,21	-0,21	0,18	-0,11	-0,18	-0,27	0,24	-0,26	0,25	-0,01	-0,05	0,13	-0,06	-0,03	-0,12	0,00	0,09	1,00	0,09	1,00

Appendix B – DEA Results for 2017-2018

Premier League 2018			Bundesliga 2018			Ligue 1 2018			La Liga 2018			Liga NOS 2018			Serie A 2018		
%NPMin	**VarPts**	**Effic**	**%NPMin**	**VarPts**	**Effic**	**%NPMin**	**VarPts**	**Effic**	**%NPMin**	**VarPts**	**Effic**	**%NPMin**	**VarPts**	**Effic**	**%NPMin**	**VarPts**	**Effic**
3,6%	-2	1	17,7%	20	1	21,4%	-5	1	5,8%	1	1	26,3%	12	1	32,2%	10	1
34,9%	22	1	20,8%	-7	0,85	22,5%	6	1	43,5%	27	1	49,0%	21	1	6,0%	5	1
33,8%	14	0,72	23,2%	2	0,76	44,9%	15	1	9,9%	-17	0,58	22,6%	-1	1	17,4%	-3	0,35
31,4%	12	0,69	25,5%	-27	0,69	37,6%	12	1,00	60,0%	21	0,58	38,7%	6	0,64	36,3%	6	0,31
21,5%	3	0,47	25,7%	-6	0,69	39,7%	11	0,88	21,6%	4	0,47	36,3%	-17	0,62	20,4%	-3	0,30
23,5%	3	0,43	25,9%	2	0,68	25,1%	-15	0,85	20,2%	3	0,43	37,7%	2	0,62	21,8%	2	0,28
11,9%	-9	0,30	26,3%	14	0,67	42,1%	11	0,83	17,8%	-20	0,32	36,7%	-6	0,62	22,1%	4	0,27
28,5%	1	0,26	26,7%	3	0,66	29,9%	-7	0,71	53,5%	8	0,30	38,2%	-19	0,59	31,8%	-12	0,19
16,5%	-23	0,21	28,3%	-14	0,62	33,1%	1	0,67	24,2%	-15	0,24	45,2%	9	0,56	33,6%	-2	0,18
19,4%	-12	0,18	30,1%	-9	0,59	33,3%	-24	0,64	35,2%	-6	0,16	43,3%	2	0,54	36,4%	1	0,17
28,1%	-1	0,17	31,3%	-7	0,56	36,6%	1	0,60	36,1%	-7	0,16	44,2%	-6	0,51	36,4%	-8	0,17
21,6%	-8	0,16	34,0%	-1	0,52	39,2%	-3	0,55	36,4%	-7	0,16	52,0%	8	0,48	41,7%	-10	0,14
26,9%	-10	0,13	35,6%	-3	0,50	53,6%	8	0,51	40,5%	-14	0,14	49,3%	1	0,47	46,5%	-5	0,13
27,1%	-3	0,13	36,4%	-12	0,49	45,8%	-4	0,47	43,0%	-3	0,13	63,8%	-3	0,35	46,6%	1	0,13
35,1%	-14	0,10	40,0%	-4	0,44	50,4%	5	0,44	49,9%	-8	0,12	64,6%	-8	0,35	46,7%	5	0,13
37,9%	-11	0,09	43,6%	7	0,41	49,0%	-17	0,44	52,9%	-26	0,11	79,9%	-1	0,28	51,5%	1	0,12
53,4%	-12	0,07	-	-	-	56,3%	-8	0,38	57,8%	-17	0,10	-	-	-	64,4%	-3	0,09

Appendix C – DEA Results for 2016-2017

Premier League 2017			Bundesliga 2017			Ligue 1 2017			La Liga 2017			Liga NOS 2017			Serie A 2017		
%NPMin	**VarPts**	**Effic**	**%NPMin**	**VarPts**	**Effic**	**%NPMin**	**VarPts**	**Effic**	**%NPMin**	**VarPts**	**Effic**	**%NPMin**	**VarPts**	**Effic**	**%NPMin**	**VarPts**	**Effic**
31,5%	43	1	8,5%	-1	1	34,6%	30	1	7,2%	3	1	36,8%	15	1	43,9%	27	1
14,0%	16	1	32,0%	25	1	19,3%	2	1	22,9%	16	1	15,1%	-6	1	26,1%	16	1
14,9%	-37	0,94	9,1%	-6	0,93	15,7%	-9	1	52,5%	20	1	49,8%	22	1	14,2%	4	1
16,3%	4	0,86	20,5%	6	0,72	26,8%	-2	0,67	11,8%	-10	0,61	20,9%	-16	0,72	19,2%	-7	0,74
17,0%	4	0,82	19,3%	-10	0,44	35,7%	9	0,65	13,6%	1	0,53	31,0%	-1	0,65	22,7%	0	0,63
21,3%	12	0,66	36,4%	6	0,41	24,6%	-17	0,64	43,9%	11	0,38	38,5%	3	0,63	25,0%	-4	0,57
24,3%	3	0,58	23,4%	0	0,40	25,7%	-10	0,61	19,3%	-1	0,37	33,8%	-4	0,51	30,8%	6	0,53
25,2%	2	0,56	23,8%	-19	0,36	33,8%	3	0,59	65,5%	13	0,29	64,0%	10	0,49	33,0%	7	0,52
25,3%	-17	0,56	23,9%	-14	0,36	27,8%	-14	0,56	26,5%	-15	0,27	64,7%	8	0,46	27,4%	-5	0,52
26,8%	-6	0,52	44,6%	7	0,35	39,1%	4	0,52	27,0%	3	0,27	34,9%	-22	0,43	33,0%	6	0,49
27,3%	16	0,51	24,3%	-8	0,35	43,4%	6	0,50	31,3%	-5	0,23	56,7%	2	0,41	29,2%	-15	0,49
27,7%	-17	0,51	28,1%	-13	0,30	58,1%	15	0,45	46,5%	-6	0,15	61,3%	-5	0,26	40,6%	8	0,45
29,0%	-7	0,48	35,1%	-3	0,24	35,4%	-16	0,44	47,3%	-6	0,15	57,7%	-17	0,26	34,8%	-14	0,41
32,8%	14	0,43	39,4%	-9	0,22	59,2%	14	0,44	49,6%	-2	0,14	71,3%	-3	0,26	40,0%	-10	0,36
33,7%	-1	0,42	42,5%	-8	0,20	40,2%	-10	0,39	50,7%	2	0,14	61,2%	-13	0,25	40,9%	-1	0,35
33,9%	-5	0,41	49,6%	-13	0,17	44,6%	-4	0,39	51,8%	-8	0,14	72,2%	-9	0,21	41,0%	-13	0,35
37,3%	-15	0,38				44,2%	-8	0,36	74,7%	-19	0,10				56,3%	8	0,32