NATIONAL TECHNICAL UNIVERSITY OF ATHENS
SCHOOL OF APPLIED MATHEMATICAL AND PHYSICAL STUDIES
DEPARTMENT OF MATHEMATICS

INTERDISCIPLINARY POSTGRADUATE SPECIALIZATION PROGRAMME
“MATHEMATICAL MODELING IN MODERN TECHNOLOGIES AND FINANCIAL ENGINEERING”

Introducing Financial Derivatives into Football Transfer Markets

MASTER THESIS

Stavros Vatikiotis

Supervisors : Dr. Athanasios Triantafyllou
Lecturer, University of Essex

Dr. Apostolos Mintzelas
KPMG Greece

Athens, June 2019

NATIONAL TECHNICAL UNIVERSITY OF ATHENS
SCHOOL OF APPLIED MATHEMATICAL AND PHYSICAL STUDIES
DEPARTMENT OF MATHEMATICS

INTERDISCIPLINARY POSTGRADUATE SPECIALIZATION PROGRAMME
“MATHEMATICAL MODELING IN MODERN TECHNOLOGIES AND FINANCIAL ENGINEERING”

Introducing Financial Derivatives into Football Transfer Markets

Stavros Vatikiotis

Supervisors : Dr. Athanasios Triantafyllou, Lecturer, University of Essex
Dr. Apostolos Mintzelas

Approved by the three-member committee

....	....	....
Dr. Athanasios Triantafyllou	Dr. Apostolos Mintzelas	Dr. Antonios Papapantoleon
Lecturer, University of Essex	KPMG Greece	Assistant Professor, NTUA

Athens, June 2019

....

Stavros Vatikiotis

MSc Mathematical Modeling in Modern Technologies and Financial Engineering, NTUA

NATIONAL TECHNICAL UNIVERSITY OF ATHENS
SCHOOL OF APPLIED MATHEMATICAL AND PHYSICAL STUDIES
DEPARTMENT OF MATHEMATICS

INTERDISCIPLINARY POSTGRADUATE SPECIALIZATION PROGRAMME
"MATHEMATICAL MODELING IN MODERN TECHNOLOGIES AND FINANCIAL ENGINEERING"

Introducing Financial Derivatives in Football Transfer Markets

MASTER THESIS

Stavros Vatikiotis

Supervisors : Dr. Athanasios Triantafyllou, Lecturer, University of Essex
Dr. Apostolos Mintzelas

Athens, June 2019

NATIONAL TECHNICAL UNIVERSITY OF ATHENS
SCHOOL OF APPLIED MATHEMATICAL AND PHYSICAL STUDIES
DEPARTMENT OF MATHEMATICS

INTERDISCIPLINARY POSTGRADUATE SPECIALIZATION PROGRAMME
“MATHEMATICAL MODELING IN MODERN TECHNOLOGIES AND FINANCIAL ENGINEERING”

Introducing Financial Derivatives in Football Transfer Markets

Stavros Vatikiotis

Supervisors : Dr. Athanasios Triantafyllou, Lecturer, University of Essex
Dr. Apostolos Mintzelas

Approved by the three-member committee

....	....	....
Dr. Athanasios Triantafyllou Lecturer, University of Essex	Dr. Apostolos Mintzelas	Antonios Papapantoleon Assistant Professor, NTUA

Athens, June 2019

....

Stavros Vatikiotis

MSc Mathematical Modeling in Modern Technologies and Financial Engineering, NTUA

Περίληψη

Κατά τη διάρκεια των τελευταίων ετών, η διαρκώς αναπτυσσόμενη βιομηχανία του ποδοσφαίρου βιώνει ένα φαινόμενο συνεχούς αύξησης του κόστους για τις μεταγραφές ποδοσφαιριστών. Η εν λόγω αύξηση αποδίδεται κυρίως στην οικονομική μεγέθυνση των ποδοσφαιρικών οντοτήτων (διοργανώσεις, σωματεία, παίκτες) μέσω εσόδων χορηγιών, στην είσοδο νέων επενδυτικών σχημάτων, αλλά και την εισχώρηση νέων ποδοσφαιρικών αγορών. Η επιβολή νέων κανόνων οικονομικής διαχείρισης από την Ευρωπαϊκή Ομοσπονδία, που αποσκοπούν στην εξυγίανση των σωματείων, επιβάλλει νέα αυστηρά πλαίσια στα οποία πρέπει να υπάρξει άμεση προσαρμογή.

Σκοπός της παρούσας διπλωματικής εργασίας είναι ο σχεδιασμός ενός νέου χρηματοοικονομικού προϊόντος βασιζόμενου στη θεωρία των Χρηματοοικονομικών Παραγώγων, το οποίο δρα συμπληρωματικά στους ήδη υπάρχοντες τρόπους μετακινήσεων και παρέχει την επιλογή στον αγοραστή να αποφασίσει εάν θα προχωρήσει στην επιθυμητή επένδυση σε ορίζοντα μιας αγωνιστικής περιόδου.

Έπειτα από τη συλλογή δεδομένων μεταβολών της αγοραστικής αξίας παικτών παγκοσμίως, όπως επίσης και ήδη ολοκληρωμένων μεταγραφών, δημιουργήθηκαν δύο μοντέλα, χρησιμοποιούμενα για τη μετατροπή αγοραστικών αξιών του σήμερα σε πιθανές μεταγραφικές τιμές στο τέλος της αγωνιστικής περιόδου. Το πρώτο δημιουργηθέν μοντέλο αποτελεί βελτίωση ήδη κατασκευασμένων μοντέλων της βιβλιογραφίας, εισάγοντας εξειδικευμένες παραμέτρους και κατ' επέκτασιν, μειώνοντας το σφάλμα προσαρμογής τους. Αντίστοιχα, το δεύτερο μοντέλο που κατασκευάστηκε, συστήνεται πρώτη φορά στη βιβλιογραφία και δεν εντοπίστηκε παρόμοιο κατά την εξέτασή της.

Εν κατακλείδι, σκιαγραφόνται τα αναγκαία επόμενα ερευνητικά βήματα, με απώτερο σκοπό την άμεση εισαγωγή του προτεινόμενου χρηματοοικονομικού προϊόντος στις ποδοσφαιρικές αγορές μεταγραφών.

Λέξεις Κλειδιά: Χρηματοοικονομικά Παράγωγα, Δικαίωμα Προαίρεσης, Μεταγραφές Ποδοσφαίρου, Αγοραστική Αξία, Τιμή Μεταγραφής, Γραμμικά Μοντέλα

Abstract

Over the last few years, the rapidly growing football industry is experiencing a continuous increase regarding the fees required for players' transfers. This increase is mainly attributable to the economic growth of football entities (competitions, clubs, players) through sponsorship revenues, the entry of new investment funds, but also the introduction of new football markets. The imposition of new economic rules by the Union of European Football Association (UEFA), aiming at the reorganization of the clubs, sets new strict frameworks that need to be adapted immediately.

The purpose of this thesis is to design a new financial product based on the theory of financial derivatives, which acts in addition to the existing ways of transferring and provides the buyer with the option to decide, whether to make the desired investment over a certain period.

After collecting market value alteration data worldwide, as well as already completed transfers, two models have been created, which are used to convert today's market values into potential transfer fees at end of the season. The first model is an improvement of an already constructed model of the literature, introducing specialized parameters, and thus, reducing the error of the model. On the other hand, the second model produced in this work, is created for the first time and has no equivalent in the examined literature.

In conclusion, the necessary further research steps are outlined, with the ultimate goal of immediately introducing the proposed financial product in football transfer markets.

Keywords: Financial Derivatives, Options, Football Transfers, Market Value, Transfer Fee, Linear Models

Dedication

To Emiliano Raúl Sala Taffarel (1990-2019) and the unpredictability of life

Acknowledgements

I would like to express my gratitude to:

- My supervisors, Dr. Apostolos Mintzelas and Dr. Athanasios Triantafyllou for their valuable guidance and advice

- My family and friends for their continuous support and encouragement

‘All models are wrong, but some are useful.’

George E. P. Box (1919-2013)

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Football Transfer Market: Definition and Current Process

Football Transfer Market is the arena in which football players are available for transfer to clubs. It consists of a transfer list of players available for transfer, and also the money moving between clubs as they contest to purchase and sell these players. During the transfer window, clubs buy replacements for players, who have suffered injuries or strengthen their squads in preparation either for an attempt to advance in a tournament or in anticipation of an upcoming struggle against relegation. For example, in Europe, there are two transfer windows (Summer: from 1st June to 31st August and Winter: from 1st Jan to 31st Jan) and only in this period players are allowed to be registered. A transfer can take place anytime during the calendar year, but the players will be ineligible to play. [1]

Generally, the whole process is conducted by the clubs and later, the players are involved. The steps followed during a typical transfer are presented:

- A club decides which squad positions need upgrades and/or backups.
- When the club is interested in signing a player, they are obliged to talk to his current

club first. It is illegal to talk to the player directly (still done discreetly, see Ashley Cole to Chelsea (2005)).

- If the parent club permits they are allowed to contact the player and his representative (known as agents).
- The club talks with both the parent club and the player in parallel trying to lure the player and conduct profitable business.
- A transfer fee is agreed upon by both clubs, which is the compensation the parent club receives for transferring the player. This fee is based on various factors like form, potential, number of years left on contract, market value, etc.
- Meanwhile the club is in negotiation with the player's agent over his new contract. Wages, length and/or bonuses and various other clauses are decided.
- If all three parties are in agreement, the transfer takes place and the player joins the club.

There are, also, some other exotic ways for a transfer to take place:

- The parent club can put the player they want to offload on the transfer list, thus making him more appealing to the other cub, despite the decrease caused to his value.
- The release clause of the player, if it exists, can be triggered. Here, the interested club basically pays the amount to break the existing contract to the player, who then forwards it to the club. Then the player can be signed by the club. It is/can be done without the consent of the parent club.
- A player whose contract has expired becomes a 'Free Agent' and any club interested can contact him or his representatives directly and sign him without having to pay additional fees to anyone else.
- The parent club sends its player to another club just on a temporary deal (on Loan), which may include a buy clause. It is usually done with young players, who go to lower leagues/teams to gain experience.

1.2 Current Situation

In recent years, there is an evident increase in the prices, that the teams have to pay to acquire the players of their interest. On Figure 1.1 [2], we present the worldwide highest transfer fees of all time as of January 2019. As we may easily see, seven out of ten of the most expensive transfers in football history were completed on the last three years.

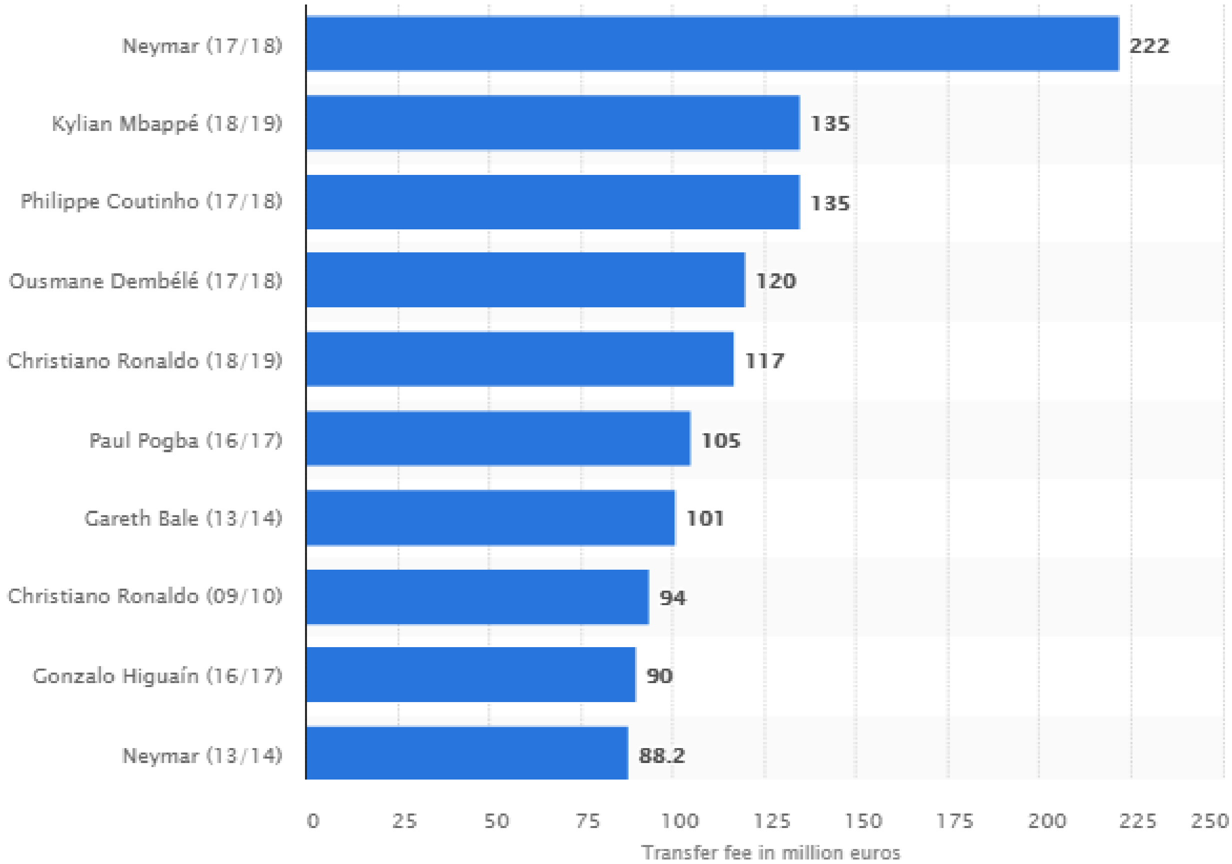

Figure 1.1: Worldwide highest transfer fees of all time as of January 2019

As FIFA's Global Transfer Market Report notes, in 2018 spending reached a new high of 7.03 billion $, 10.3 % more than in 2017. A new record was set with 16,533 international transfers, 5.6% more than in 2017, involving 14,186 players of 175 different nationalities. In addition, it is important to state that only 31 clubs spent more than 50 million $ each, but together they accounted for more than half of the 7.03 billion $ spent globally. [3]

On the following figure [4], we observe that the biggest percentage of the top-100 most expensive players is located during the past five years, and, this trend seems to be continuously increasing.

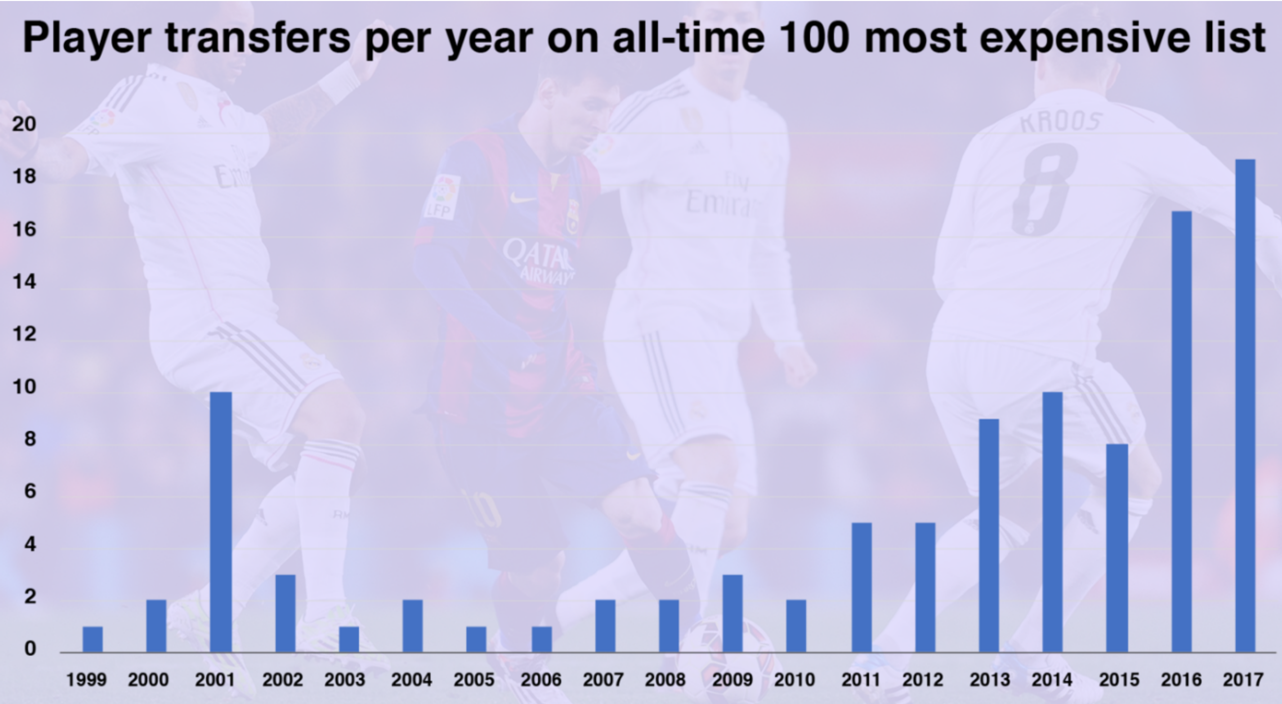

Figure 1.2: Average and largest transfer fee in recent years

This outburst can be explained as a result of several factors, such as:

- The increase in the operating revenues of football organizations through TV rights and advertising, as football is one of the most growing industries during the last 40 years.
- The increase of the financial inflows in the football field, as more investment funds emerge declaring interest to invest, either in football clubs or even in football players' rights.
- The entry of new markets in the football field, such as the Asian championships.

We present on Figure 1.3 [5] the transfer expenses categorized by the buyer's league on winter of 2016, which was the first transfer window that Chinese Super League clubs were so active. It is obvious that Chinese clubs spent more money than every European League, while the transfer fee for some players was much higher than their value at the time.

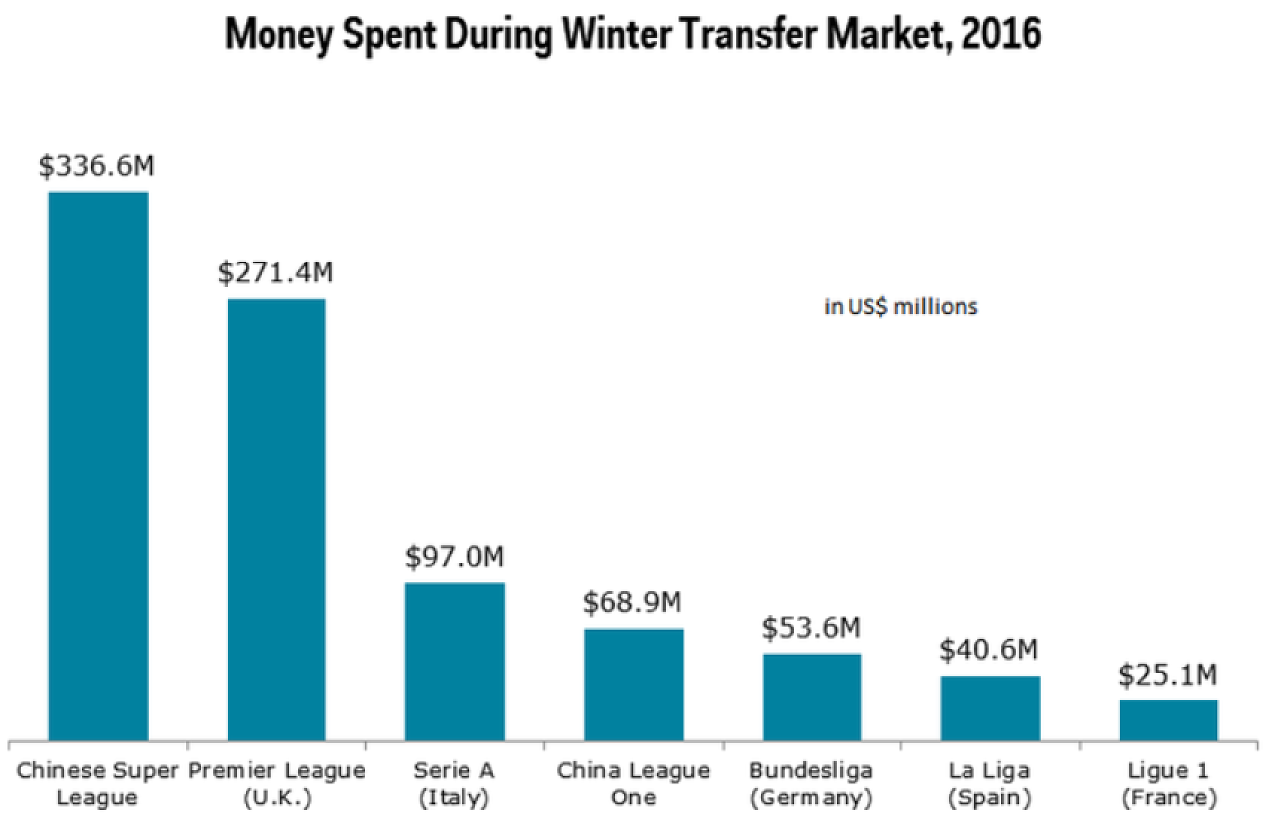

Figure 1.3: Transfer Market Expenses during winter transfer market 2016

1.3 What about Financial Fair Play?

This "high prices phenomenon" is in contrast to the Financial Fair Play (FFP) rules submitted by UEFA in 2010, which aim to control the viability of the clubs by decreasing their debt. It is important to note that one of the key aspects of FFP is limiting their external financial dependence on their major shareholders. To become more specific, if a club's owner injects money into the club through a sponsorship deal with a company to which he is related, then UEFA's competent bodies will investigate and, if necessary, adapt the calculations of the break-even result for the sponsorship revenues to the level which is appropriate ('fair value') according to market prices. Under the updated regulations, any entity that, alone or in aggregate together with other entities which are linked to the same owner or government, represent more than 30% of the club's total revenues is automatically considered a related party.

To underline how crucial the matter is, clubs that fail to comply with FFP rules are facing the following sanctions:

- warning
- reprimand
- fine
- deduction of points
- withholding of revenues from a UEFA competition
- prohibition on registering new players in UEFA competitions
- restriction on the number of players that a club may register for participation in UEFA competitions, including a financial limit on the overall aggregate cost of the employee benefits expenses of players registered on the A-list for the purposes of UEFA club competitions
- disqualification from competitions in progress and/or exclusion from future competitions
- withdrawal of a title or award

Combining the aformmentioned measures and the football transfer market inflation phenomenon, we understand, that clubs equipped with improved financial tools and better statistical resources are most likely to overcome the dangers that arise. [6]

1.4 Terminology

We will now present some important terms that will be used throughout this thesis so as to make more understandable the process followed.

- **Market Value** is the actual value of the player as it is estimated by the player's performance and achievements in the football pitch.

- **Possible Transfer Fee** , is the fee that a club would be willing to pay to acquire the player's rights, and in reverse, the fee that a club would demand to release the player. This amount depends on the player's market value, but under no circumstances should be considered as the same. For example, we know empirically that as the player's contract with a particular club approaches its expiration date, clubs will try to sell the player at a lower price because they face the risk of losing the player without receiving any compensation at all. However, this fact does not affect the player's market value which is determined by the player's achievements.

- **Performance Index** , is the rating the player gets for his performance depending on the position on the field.

- **Domestic League** , is the domestic championship where the player's club participates.

- **Champions League** ,is the Europe's biggest international competition organised by UEFA (Union of European Football Associations) and the participation of the clubs is determined by the position held in their domestic leagues. It provides not only glory for the winner but also increases the cash inflows for the clubs through ticket sales, TV advertising and bonus given by UEFA.

- **Europa League** , is the Europe's second biggest international competition, with similar benefits to the UEFA Champions League. Important addition to the competition was the opportunity for the winner of the competition to qualify automatically for next season's Champions League regardless of the position possessed in the domestic league.

- **UEFA Country Coefficient** , is used to rank the football associations of Europe, and thus determine the number of clubs from an association that will participate in the UEFA Champions League and the UEFA Europa League.

  The UEFA ranking determines the number of teams competing in the season after the next, not in the first season after the publication of the ranking. Thus, the rankings at the end of the 2017/2018 season determine the team allocation by association in the 2019/2020 (not 2018/2019) UEFA season. This is unrelated to the selection of teams which will fill each allocation through the individual association leagues and national cups (which is decided in the preceding year).

This coefficient is determined by the results of the clubs of the associations in the UEFA Champions League and the UEFA Europa League games over the past five seasons. Two points are awarded for each win by a club, and one for a draw (points are halved in the qualifying and playoff rounds). To determine a country's coefficient for a particular season, the coefficients for the last five seasons are added. Bonus points are added to the number of points scored in a season. Bonus points are allocated for: clubs that reach the quarter-finals, semi-finals, or final of either the UEFA Champions League or the UEFA Europa League (1 bonus point), clubs that qualify for the group stage of the Champions League (4 bonus points) and, finally, clubs that qualify for the round of 16 of the Champions League (5 bonus points).

- **UEFA Club Coefficient** , The club coefficient rankings are determined by the results of clubs in the UEFA Champions League and the UEFA Europa League over the previous five seasons, as well as by the coefficient of the clubs' association. The club coefficient is the sum of the points earned by the club over the five seasons, but no less than 20 % of the club's association coefficient. Prior to the 2018 club rankings, teams received the sum of their points earned over the last five seasons plus 20% of the club's association coefficient.

  The clubs receive two points for a win, one point for a draw, and no points for a defeat in games of the main stages of the Champions League and the Europa League. Results determined after extra-time are included in this method, however results determined after penalty shoot-outs are not (the result is considered a draw). Bonus points for entering the Europa League group stage are not additional to win/draw points; they provide a minimum points allowance for participating clubs, whereas bonus points for entering the Champions League group stage (and those for qualifying to the knockout stage) are additional to win/draw points. [1]

Chapter 2

The proposed financial product

In this section, we present the main idea of this thesis, as well as the possibilities that arise.

2.1 Main Idea

After examining the current situation as well as the clubs' vulnerabilities in the Introduction, we have to create a financial instrument designed for football clubs to use, in order to achieve more efficient deals. It is indispensable for our proposed financial product to:

- provide time to the buying club to thoroughly evaluate the player
- provide time to the selling club to replace the player
- provide the option to the buying club not to acquire the player at the end of the agreed time window
- have the potential to be used in the secondary market
- set specific transfer fees today, which will be paid at the end of the time window

These characteristics may be considered as the basis of an option in Financial Derivatives. Therefore, we conclude that our proposed financial product needs to be a call option between the two clubs, with the player being the underlying asset.

2.2 Financial Derivatives Theory

Before we continue with our methodology, it is essential to present the theoretical concept of derivatives, so as to understand the mathematical relations and numerical values, we are required to estimate.

Derivatives are contracts between two parties that specify conditions (especially the dates, resulting values and definitions of the underlying variables, the parties' contractual obligations, and the notional amount) under which payments are to be made between the parties. The assets include commodities, stocks, bonds, interest rates and currencies, but they can also be other derivatives, which adds another layer of complexity to proper valuation.

From the economic point of view, financial derivatives are cash flows, that are conditioned stochastically and discounted to present value. The market risk inherent in the underlying asset is attached to the financial derivative through contractual agreements and hence can be traded separately. That contractual freedom allows derivative designers to modify the participation in the performance of the underlying asset almost arbitrarily.

Derivatives may broadly be categorized as "lock" or "option" products. Lock products (such as swaps, futures, or forwards) obligate the contractual parties to the terms over the life of the contract. Option products (such as interest rate swaps) provide the buyer the right, but not the obligation to enter the contract under the terms specified.

Derivatives can be used either for risk management (i.e. to "hedge" by providing offsetting compensation in case of an undesired event, a kind of "insurance") or for speculation (i.e. making a financial "bet").

Options: Usage and Basic Terms

There are two types of options. A call option gives the holder the right to buy the underlying asset by a certain date for a certain price. A put option gives the holder the right to sell the underlying asset by a certain date for a certain price. The price in the contract is known as the exercise price or strike price; the date in the contract is known as the expiration date or maturity. There two main types of options are; American options that can be exercised at any time up to the expiration date and European options can be exercised only on the expiration date itself. It should be emphasized that an option gives the holder the right to do something. Buyers are referred to as having long positions; sellers are referred to as having short positions. [7] On the following graph, we present the possible profit from each state in an option.

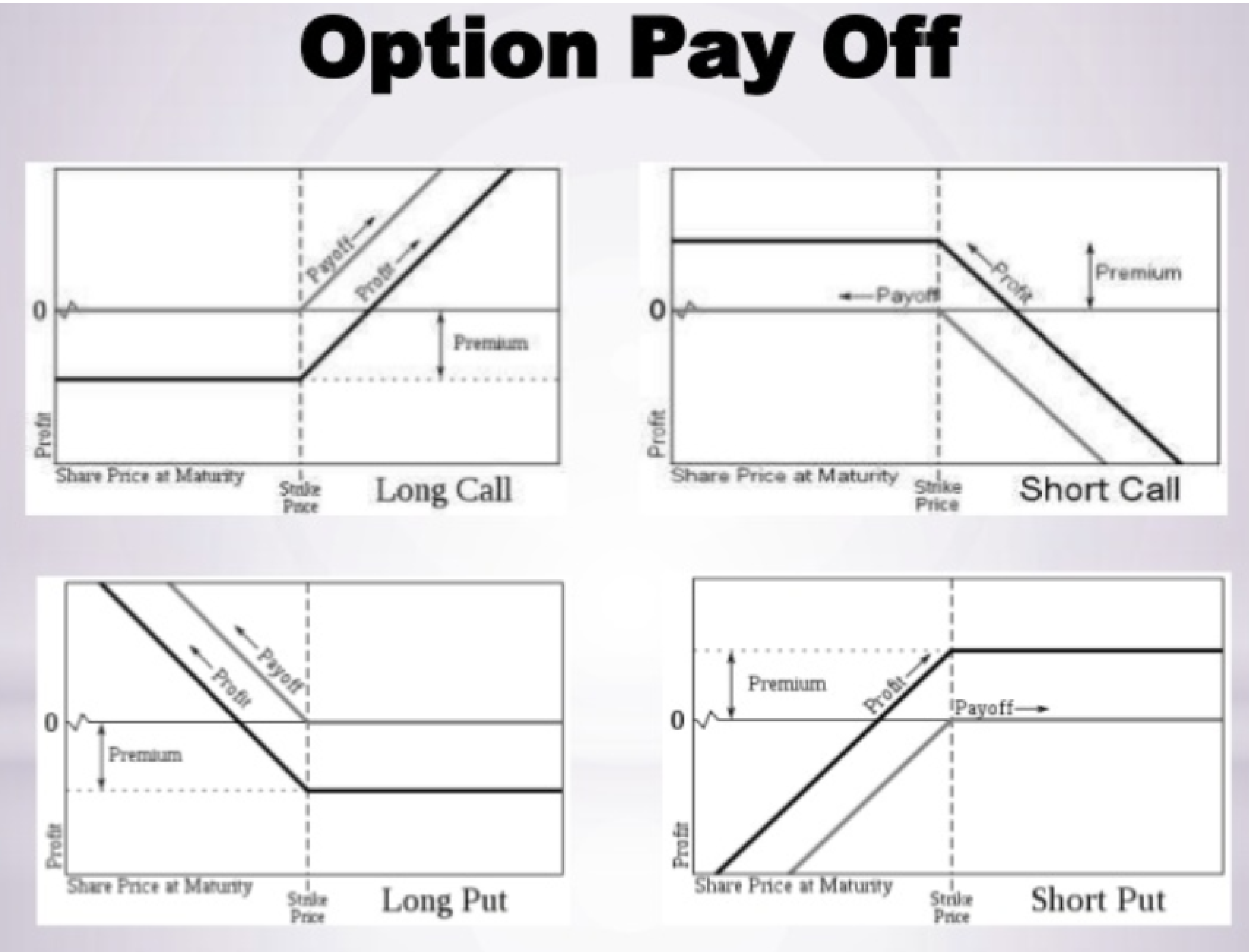

Figure 2.1: Option Pay off categorised by type

2.3 Methodology

It is, now, clear that the proposed financial product is a call option in which the buying club is long and the selling club is short. We, also, assume that the option is signed on the summer transfer window, while the maturity date is set at the end of the season (for example, on the 30th of June). Therefore, the key objective of this work is to provide an accurate estimation regarding the player's possible transfer fee at the end of the season. As we have already mentioned, there is a bonding relationship between a player's possible transfer fee and the corresponding market value. Thus, we will firstly develop a model, which connects the player's current market value to the market value at the end of the season, in accordance to the factors that are commonly used to alter this value. The second step would be the development of another model, which transforms the market value to the possible transfer fee (Figure 2.2).

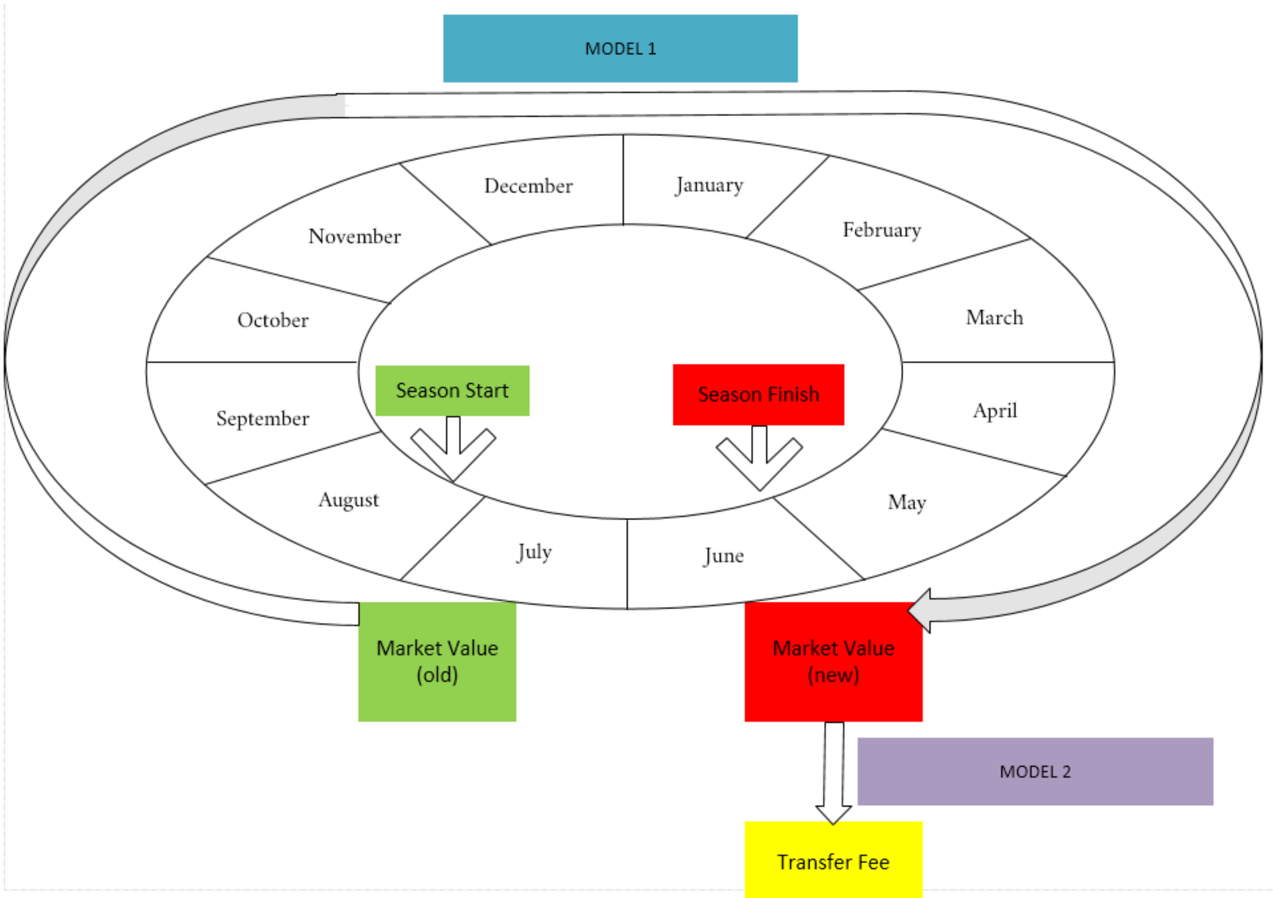

Figure 2.2: Our Methodology Diagram

In order to clarify the concept, it is of great use to provide a possible example of the product's application in a real life scenario.

Suppose that Club A is interested to buy football player XYZ from Club B. XYZ is a young talented striker, who has just started his professional career and the team's scouting report is encouraging regarding his progress. On the other hand, the young player has not proved yet his consistency, as far as performance is concerned, and the team manager is reluctant to invest valuable transfer capital on buying XYZ. However, the most important problem for the transfer's completion is that currently, club A already possesses four other footballers that play in the same position, making very difficult for the young player to get minutes in action, if signed this summer.

First Scenario: Transfer with current market tools

Supposing that Club A is not interested in risking losing the player and by using the current tools of football transfer market, Club A has two choices; they either sign XYZ and keep him in their roster, risking to impede his progress, since he will not play enough minutes to develop, or sign the player and send him on loan (most likely to Club B). The second strategy is a very common phenomenon nowadays, as there are no direct obstacles to the player's development. Having this in mind, Club A buys the player by paying the current price set by Club B, while Club B is allowed to keep the player in their for the next season. Unfortunately, though, XYZ gets severely injured and according to the medical staff, he will never play to his real potential. As a result, his market value is deteriorating rapidly, while Club A invested an important share of their transfer budget on a player, who will never get to the level to play for them. And this is only one bad hypothesis, among any possible worst case scenarios.

Second Scenario: Transfer using the proposed product

In this scenario, the starting point remains the same but Club A has a more flexible way to arrange the transfer. A meeting between the two clubs Sporting Directors is held, where Club A proposes the following: 'We are interested in XYZ but we would like to buy him next summer at a price we would agree right now. Your club should not worry for this season, as he will keep on playing for your team. Just to clarify, if at the end of this season we are not interested any more, we have the option of not completing the transfer. Finally, you will now receive some compensation for signing this deal.' After thinking for a while, the Sporting Director of Club B responds 'All this sounds very interesting, but what is the amount of money we will receive now, and at the end of the season?'

The estimation of these two amounts is one of the main objectives of this work!

2.4 Existing Literature

By reviewing the existing literature, we have come to the conclusion that so far, there have been two separate approaches regarding market value alteration and transfer fee estimation; stochastic and regression.

2.4.1 Stochastic Approaches

R.Tunaru et Al. (2005) in their work "An option pricing framework for valuation of football players" used a Geometric Brownian Motion for each player's index points and then, transformed the estimated index points into cash. It is important to note that there has been assumed a distinguished valuation between the club owning the player and outside clubs. Another interesting contribution is the inclusion of a Poisson process to model a possible abrupt increase (performing extremely well) or decrease (injury). [8]

2.4.2 Regression Approaches

Herm et Al. (2014) estimated the player's current market value of 338 active Bundesliga athletes, using two main categories; talent (attributes of age, precision, scoring assertion and flexibility) and external factors (focusing mainly on player's popularity with features such as public attention, player's agent, club management and team coach). As players of different positions cannot be criticized based on the same criteria (for example, a very talented defender will not score as many goals as an attacker, and vice versa), the talent values have been corrected using the mean category values of each position. [9]

Furthermore, Miao He, Ricardo Cachucho and Arno Knobbe on their conference paper "Football player's performance and market value" (2015) have used a dataset consisting of 381 Spanish La Liga players. That work's goal was to provide a strong relationship between a player's market value and his corresponding performance. Moreover, a special model fitted for attacking players has been produced using performance stats (goals scored, assists provided, fouls committed etc.).[10]

Additionally, O. Muller et Al. (2017) estimated the players' possible transfer fee using a 10,350 observation dataset of only outfield players (there were no goalkeepers used for the model). The features used were focusing on the players' characteristics (age, height), statistic performance of the player (minutes played, goals, passes, cards received etc.) and popularity measures such as wikipedia page views, youtube videos and reddit posts. [11] This model seems to be more personalized compared to the model created in Herm et Al. (2014), as the popularity data used were directly derived by each player and there was no need for correction values through the sample itself. It is also important to note that Muller's observations come from the top-5 European Leagues in contrast to the other approaches that use only German Bundesliga and Spanish La Liga players respectively.

However, it is important to note that in neither of the aforementioned models, there has been a distinction between the market value of a player and the possible transfer fee. As mentioned above, in Introduction's Terminology, this distinction is of high importance, as a major feature of the way football transfer markets operate is the player's current contract's expiration date, which is not reflected to the player's personal market value.

2.5 Created Opportunities

After presenting these scenarios, it is rational to believe that this financial product can be of much help sealing transfer deals in the future. Here are the main reasons supporting this conviction:

- **This financial option can be used as an asset for another deal** , as the holder of the option could include it as a part of another transfer to lure the corresponding selling club. As we have already seen, from the financial derivatives theory, options can be valued at any time, until the maturity date using the Black Scholes formula. Apart from the financial aspect, though, the underlying asset i.e. the player, might be of some interest for the second club and could be considered as a bargain at a certain price, without having to pay the premium price.

- **The long position club may buy the player for a lower price** , as in nowadays' transfer market, a good year could possibly launch a player's price very high. Additionally, a very important thought about this argument is that the demanded fee is adapted to the buyer's financial equilibrium between sales and buys. For example, if a club manages to have a massive profit throughout a transfer window and is interested into buying a player, the selling club will surely ask for more money than the player's fair price.

- **The player realizes that the transfer is not yet completed** , so he is obligated to keep working hard to prove he is adept enough to get it. It is also high likely, that the player will maintain his playing time and thus, having no loss in aspects of match fitness.

- **The short position club can use the player for at least one more year** , so there is no need of buying another player to replace him immediately. Additionally, the manager's squad plans are not directly modified, and of course, the fans could not be moaning for losing a player - at least- yet.

- **The club in the long position is not obligated to buy the player** , as the main advantage of the financial options as an instrument, is that they provide a choice and not an obligation. There underlies a frequent situation in football. Sacking a manager at the end of the season and hiring a new one, while having players who do not fit his playing style. Supposing that a huge fee has been paid for these players, the club directors fall into a dilemma; keeping the player but not completing the deal of a -probably very skillful- manager or selling the player and suffer a financial loss.

- **The short position club can instantly improve their current liquidity status** , as the premium is instantly paid and could probably be used to buy another player in positions that the team suffers, or repaying any debt created in the past, but without losing an important member of the squad.

- **Hedging** , as it is a very common phenomenon for clubs to negotiate a player's contract renewal very close to the expiration date. We should keep in mind that six months before the expiration date, a player is considered a free agent and may leave the club without it receiving any compensation at all. When the club tries to replace this player by signing another with similar characteristics, such as position, selling clubs tend to increase their demands, because they conceive this situation as an opportunity to capitalize on the emergency. However, using the proposed financial derivative, clubs have the option to replace the player that is not willing to renew his contract, before this situation is widely known. Furthermore, in case the player changes his mind and decides to remain at the club, the club has the option not to exercise the derivative.

2.6 Two Real Life Cases

Finally, we present two cases obtained from the actual football transfer market to support the aforementioned theoretical points.

2.6.1 The Naby Keita Transfer

Naby Laye Keïta (born 10 February 1995) is a Guinean professional footballer, who currently plays as a central midfielder for Premier League club Liverpool and captains the Guinea national team.[1] The point of interest in our case is the way he was transferred from RB Leipzig to Liverpool.

On 28 August 2017, a deal was struck for Keïta to join Liverpool on 1 July 2018 after the English club triggered his £48 million release clause in addition to paying an undisclosed premium. There would be no premium (£48 million total) if Leipzig did not qualify for European football, £4.75 million (£52.75 million total) if they qualified for the Europa League and £11 million (£59 million total) if they finished in the Champions League spots. Leipzig finished 6th in the Bundesliga, qualifying for the Europa League. [12]

For the reasons already mentioned on paragraphs 2.2 and 2.3, we conclude that this transfer deal has been very risky from Liverpool's perspective. Indicatively, we may repeat some scenarios, which would make this deal unsuccesful:

- A severe injury happens to Naby, which will never let him play to his actual potential.
- Naby gets more excited about his transfer than he should have been, and stops performing the way he should.
- Naby's previous performances have been just a "firework" and it is finally proved that he is not "Liverpool material".

Even though, none of the above hypotheses were confirmed in Keita's season 2017/2018, we understand that all this scepticism is justified, as all these scenarios could have occurred in the unpredictable world of football. According to transfermarkt.de, during 2017/2017 season, Naby was ruled out for a month due to injury problems, while in 2017/2018 campaign, he has been ruled out from 3 RB Leipzig matches for the same reason. [13]

Finally, it is worth mentioning that there were probably included some clauses in the transfer deal (apart from the bonus depending on the RB Leipzig 's final position in the German Bundesliga 2017/2018), but exact information on the matter have not been disclosed. However, the most precise clause that might have been included should provide the option for Liverpool, to determine if they want to finalise the deal on 2018, which would make this transfer almost equal to our proposed financial product.

2.6.2 The FC Barcelona Transfer Ban

Exactly five years ago, FIFA confirmed an unprecedented transfer ban on Barcelona, ruling that the Catalan club would not be able to sign new players during the next two transfer windows. The Blaugrana not only survived, they thrived - back-to-back domestic doubles and a Champions League title represent the most successful two-year stretch in the club's storied history. But the trophy haul does not tell the entire story. The ban may have irreparably stalled the careers of Arda Turan and Aleix Vidal, two 2015 summer signings that were forced to wait until January to make their debuts. The resulting lack of depth nearly caused a historic collapse at the end of last season, as Luis Enrique's distrust of his bench wore down the starting XI. Still, with Real Madrid and Atletico Madrid facing transfer moratoriums of their own, Barcelona's success could serve as a blueprint for their two rivals. FIFA handed down the 14-month transfer ban in April, 2014, when Barcelona were found to have violated Article 19 of the FIFA Regulations on the Status and Transfer of Players by signing international players under the age of 18. An appeal pushed the punishment to 2015, allowing the club to sign Luis Suarez, Ivan Rakitic, Jeremy Mathieu, Claudio Bravo and Marc-Andre ter Stegen. The January 2015 window was inconsequential - the Blaugrana rarely conduct much mid-season business, and the club went on to win the LaLiga-Copa del Rey-Champions League treble. Xavi and Pedro left the Nou Camp on summer 2015, and Luis Enrique brought in Turan and Vidal to bolster a very thin bench. When the duo finally debuted, six months after signing, they struggled from the outset. The Turkish playmaker never found his position in Luis Enrique's preferred 4-3-3 and started just one of the club's last 10 matches. Vidal was even less of a factor - the full-back's last appearance was against Getafe on March 12, and he was dropped from the team entirely for a stretch in early April. Their ineffectiveness and a spree of untimely injuries forced the Asturian coach to lean heavily on his starters. Suarez played every minute of every Champions League match and 36 league games, Lionel Messi sat out 45 minutes in all of 2016, and Neymar was substituted once all season. Those tired legs caused an April freefall that nearly cost Barcelona their season - three-straight league losses allowed Los Blancos and Atletico to pull within a single point, and the Rojiblancos knocked the Catalans out of the Champions League. But the Blaugrana finished strongly, and their LaLiga title ensures that the transfer ban will be a small footnote in the legacy of the MSN era. [14]

From this Marca aftermath report, we understand that this ban has immensely affected the club's performance and it would be fair to conclude that this sudden crisis has not been effectively dealt with. However, if all the aforementioned are not problematic enough we may present one more aspect to the matter. The next transfer window where FC Barcelona were eligible to register new players was summer 2016. The signings of that transfer window were [13]:

- Andre Gomes from Valencia CF for 37,00 Mill. euros.
- Paco Alcácer from Valencia CF for 30,00 Mill. euros.
- Samuel Umtiti from Olympique Lyon for 25,00 Mill. euros.
- Lucas Digne from Paris Saint-Germain for 16,50 Mill. euros.
- Jasper Cillessen from Ajax Amsterdam for 13,00 Mill. euros.
- Denis Suarez from Villarreal CF for 3,25 Mill. euros.

All these transfer moves sum up to expenditures of 124,75 millions of euros. From the aforementioned players, only two have remained in FC Barcelona squad: Samuel Umtiti, who despite having won the FIFA World Cup 2018 playing for France, struggled with major injury problems during 2018/2019 and heis reportedly on the verge of leaving the club, and Jasper Cilessen who is the second in hierarchy goalkeeper, having played in only 31 matches (23 of them in Copa del Rey). The rest of the players have been either loaned several times or sold, but in no means approaching the investment made. [13]

From these observations, we conclude that not only the club has not made the optimal choices to deal with this crisis throughout the transfer ban period, but also, that other clubs felt that this situation could have easily been exploited (as they did), considering that due to the FC Barcelona not effective transfer budget allocation, there were money resources available.

2.6.3 Case Studies Output

From both of these stories that recently happened, we receive some precious feedback. Firstly, we feel that clubs are not provided with the appropriate tools to assess the underlying risk of their potential investments. Secondly, we understand that not only clubs lack experience in managing sudden emergency situations in terms of transfer deals or employing different ways of signing a player, but most worryingly, they lack the necessary financial instruments.

As more is yet to come, it will be very interesting to observe how clubs will face these situations from now on, considering that there is already a precedent. For example, the most recent team set to be handed a transfer ban sanction [15], Chelsea FC has already paid Borussia Dortmund 64,00 millions of euros for 20 year old, American, Christian Pulisic, while the player's contract with the German club was due to expire on June 2020, meaning that from January 2020, the player would be considered as a free agent.[13]

Chapter 3

MaVAM: A Market Value Alteration Model

In this section, we present the first model developed for this work which estimates the alteration of the player's market value during a single football season (Figure 3.1).

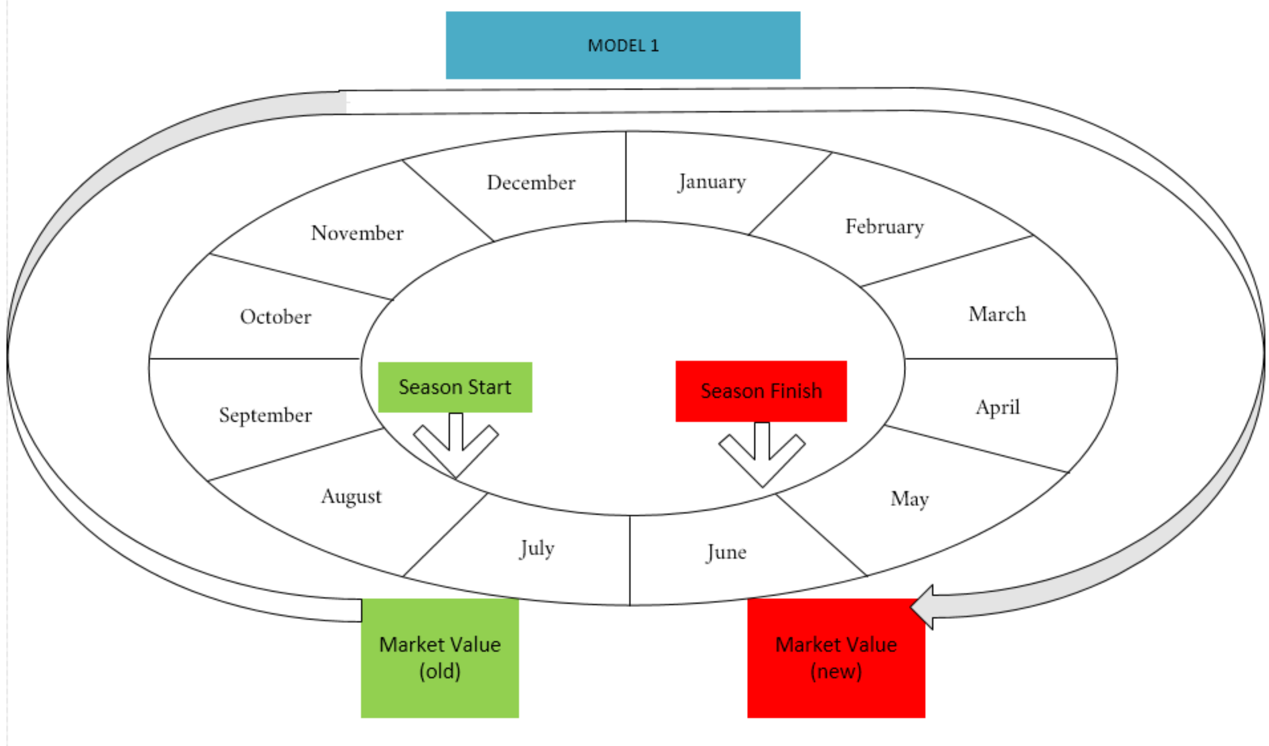

Figure 3.1: Our Methodology Diagram

3.1 Novelties of MaVAM

Since we have studied the existing literature on football players' market value, we should proceed to justify why MaVAM makes a real difference on the subject. When starting to create a model, one has to question himself, which variables have important impact on the model's response. In a similar fashion, we had to position ourselves on the fans' state of mind and understand the criteria used, in order to determine a player's new market value. Through this process, it became obvious that apart from all the factors that have already been used in previous works, players' market values are highly affected by some additional parameters, which had to be employed on MaVAM.

To begin with, MaVAM separates a player's international performance from the corresponding domestic. As we will show on the next steps of this work, international minutes played on Champions League or/and Europa League, as well as the player's performance during the competition are important aspects, which should not be neglected.

Furthermore, we realised that the differentiation between a player's domestic and international stats was encouraging but not absolutely satisfying. There has to be a distinction among the domestic Leagues that a player takes part in. It is rational to assume that, for example, a player competing in English Premier League may improve his market value more rapidly than another who plays for a club, which does not participate in one of the Top-5 European football Leagues. Apart from that, even between the Top-5 football Leagues, there has to be a different Coefficient included. The most appropriate measure for this work is considered to be the UEFA Country Coefficient that provides the every year's rankings for each League, depending on the countries' clubs progress in the international competitions.

Last but not least, the final major achievement of MaVAM is that for the first time, a distinction between clubs is used. Empirically, we understand that a player who manages to play for a top team, will most likely observe an important increase on his market value. Thus, we will also use the UEFA Club Coefficient, that shows a club's progress in the international competitions.

It is important to note, that UEFA Coefficients were used as an effective and trustworthy measure, because of their ability to balance between consecutive success, which implies that a club is highly consistent, and a sole yet remarkable season's campaign.

3.2 The Dataset Used

We shall now proceed to the explanation of the dataset used for the model production. The data gathered for approximately 2000 unique players (5475 observations in total, as seasons 2016/2017, 2017/2018 and 2018/2019 were used in producing the model) were the following:

- **Age 2018:** age of the player on the summer of 2018
- **Age 2017:** age of the player on the summer of 2017
- **Age 2016:** age of the player on the summer of 2016
- **Domestic League Minutes in Play during season 2018-2019:** The minutes the player has participated in Domestic League matches for his team during season 2018-2019
- **Domestic League Minutes in Play during season 2017-2018:** The minutes the player has participated in Domestic League matches for his team during season 2017-2018
- **Domestic League Minutes in Play during season 2016-2017:** The minutes the player has participated in Domestic League matches for his team during season 2016-2017
- **Domestic League Performance Index during season 2018-2019:** The player's evaluation in Domestic League matches during season 2018-2019
- **Domestic League Performance Index during season 2017-2018:** The player's evaluation in Domestic League matches during season 2017-2018
- **Domestic League Performance Index during season 2016-2017:** The player's evaluation in Domestic League matches during season 2016-2017
- **Champions League Minutes in Play during season 2018-2019:** The minutes the player has participated in Champions League matches for his team during season 2018-2019

- **Champions League Minutes in Play during season 2017-2018:** The minutes the player has participated in Champions League matches for his team during season 2017-2018
- **Champions League Minutes in Play during season 2016-2017:** The minutes the player has participated in Champions League matches for his team during season 2016-2017
- **Champions League Performance Index during season 2018-2019:** The player's evaluation in Champions League matches during season 2018-2019
- **Champions League Performance Index during season 2017-2018:** The player's evaluation in Champions League matches during season 2017-2018
- **Champions League Performance Index during season 2016-2017:** The player's evaluation in Champions League matches during season 2016-2017
- **Europa League Minutes in Play during season 2018-2019:** The minutes the player has participated in Europa League matches for his team during season 2018-2019
- **Europa League Minutes in Play during season 2017-2018:** The minutes the player has participated in Europa League matches for his team during season 2017-2018
- **Europa League Minutes in Play during season 2016-2017:** The minutes the player has participated in Europa League matches for his team during season 2016-2017
- **Europa League Performance Index during season 2018-2019:** The player's evaluation in Europa League matches during season 2018-2019
- **Europa League Performance Index during season 2017-2018:** The player's evaluation in Europa League matches during season 2017-2018
- **Europa League Performance Index during season 2016-2017:** The player's evaluation in Europa League matches during season 2016-2017
- **Position:** The player's main position on the pitch. The subcategories of this factor are; Goalkeeper, Defender, Midfielder, Attacker.
- **Market Value on June 2019:** The player's market value on June 2019

- **Market Value on August 2018:** The player's market value on August 2018
- **Market Value on June 2018:** The player's market value on June 2018
- **Market Value on August 2017:** The player's market value on August 2017
- **Market Value on June 2017:** The player's market value on June 2017
- **Market Value on August 2016:** The player's market value on August 2016
- **Country UEFA Coefficient for 2018-2019 season:** The league's UEFA coefficient just before 2018-2019 season starts
- **Club UEFA Coefficient for 2018-2019 season:** The club's UEFA coefficient just before 2018-2019 season starts
- **Country UEFA Coefficient for 2017-2018 season:** The league's UEFA coefficient just before 2017-2018 season starts
- **Club UEFA Coefficient for 2017-2018 season:** The club's UEFA coefficient just before 2017-2018 season starts
- **Country UEFA Coefficient for 2016-2017 season:** The league's UEFA coefficient just before 2016-2017 season starts
- **Club UEFA Coefficient for 2016-2017 season:** The club's UEFA coefficient just before 2016-2017 season starts

The above data were respectively extracted from these sources:

[1]: transfermarkt.de

[2]: whoscored.com

[3]: uefa.com

It is also important to explain the way these sources produce their data.

- **transfermarkt.de** , is Germany's leading online soccer community and one of the largest sport websites on the German language Internet and applies the judge principle. Every interested person can register for free on the transfermarkt.de community and discuss market values of thousands of soccer players playing in various countries and divisions. Commonly, a member proposes a market value for a player and provides a rationale for this number, for example, by arguing about one or more of the player's attributes to justify his or her market-value estimation. Every community member can follow the discussion threads of all the players' market values and users can contribute based on criteria that are most important to them personally. However, only a few merited community members are discussion leaders – "judges" in the language of our research. Judges have earned superior rights and make the final decisions about market-value suggestions that are posted by community members. Market values provided by transfermarkt.de have a good reputation in the sports industry and have a high economic relevance; they are used in actual transfer and salary negotiations. Also they have been used as a proxy for income or market value in scientific research (Bryson, Frick, Simmons, 2013; Franck Nuesch, 2012).

- **whoscored.com** ,is a unique website and one of the fastest growing in the sports industry, specializing in the in-depth analysis of detailed football data. Ratings are calculated based on a non public algorithm, using OPTA's statistics and are updated during each game. The Rating variable is scaled from 0-10 where 10 indicates best.

- **uefa.com** , is the official website of the "Union of European Football Associations", which provides information for all the competitions held, but more importantly on this case, the Club and Country Coefficient Rankings that are produced by a certain point system.

As already mentioned, this unique dataset consists of 5475 observations i.e. players from 11 football leagues in Europe (English Premier League, Spanish La Liga, Italian Serie A, German Bundesliga, French League Un, Portuguese Liga NOS, Dutch Eredivisie, Turkish SuperLig, English Championship, Brasileirao and Argentinian League). Its detailed composition is presented on the following tables:

***EPL**: England, **SLL**: Spain, **GBD**: Germany **ISA**: Italy, **FLU**: France, **RoW**: Rest

Players' Categorisation					
Players	16-20	21-23	24-26	27-30	31+
EPL GK	0	8	19	37	32
EPL Def	29	75	127	126	48
EPL Mid	33	62	97	101	20
EPL Att	30	55	109	88	12
EPL Total	**92**	**200**	**352**	**352**	**112**
SLL GK	4	15	17	30	19
SLL Def	27	68	93	102	27
SLL Mid	28	58	63	71	19
SLL Att	27	64	67	62	17
SSL Total	**88**	**205**	**241**	**265**	**82**
GBD GK	2	10	18	30	11
GBD Def	57	75	87	60	22
GBD Mid	46	65	74	68	10
GBD Att	52	63	81	60	16
GBD Total	**157**	**213**	**260**	**218**	**60**
ISA GK	8	12	17	18	27
ISA Def	25	87	77	66	46
ISA Mid	41	68	67	65	33
ISA Att	25	61	58	48	19
ISA Total	**99**	**228**	**219**	**197**	**125**
FLU GK	5	14	25	16	18
FLU Def	39	72	71	55	42
FLU Mid	41	76	48	40	16
FLU Att	46	58	45	38	10
FLU Total	**132**	**220**	**189**	**149**	**86**
RoW Total	**159**	**222**	**270**	**201**	**82**
TOTAL	**727**	**1288**	**1531**	**1382**	**547**

Table 3.1: Players' Categorisation by Age, Position, League

Domestic Minutes Played Mean Values					
Players	16-20	21-23	24-26	27-30	31+
EPL Total	1197,84	1618,49	1867,04	1835,3	1763,7
SLL Total	1366,96	1717,21	1851,89	1928,71	1632
GBD Total	1316,56	1599,69	1683,23	1625,41	1604,28
ISA Total	1472,47	1672,6	1785,69	2003,53	1877,17
FLU Total	1542,11	1853,50	1883,82	1934,85	1867,54
RoW Total	1541,58	1882,39	2070,48	2049,41	1921,7
Total	1406,25	1723,98	1857,03	1896,20	1777,73

Table 3.2: Players' Categorisation by Minutes Played in Domestic League

Domestic Performance Index Mean Values					
Players	16-20	21-23	24-26	27-30	31+
EPL Total	6,65	6,77	6,79	6,76	6,72
SLL Total	6,71	6,74	6,76	6,82	6,74
GBD Total	6,71	6,77	6,77	6,80	6,76
ISA Total	6,64	6,75	6,81	6,87	6,76
FLU Total	6,74	6,81	6,84	6,81	6,76
RoW Total	6,85	6,94	6,99	6,93	6,96
Total	6,72	6,80	6,83	6,83	6,78

Table 3.3: Players' Categorisation by Index Mean Values in Domestic League

Observations participating in Competitions				
Season	Domestic	Champions League	Europa League	Both
2016/2017	1779	399	333	58
2017/2018	1900	411	348	81
2018/2019	1796	496	369	75
Total	5475	1306	1050	214

Table 3.4: Players' Categorisation by Competition

From Tables 3.2, 3.3 and 3.4, we obtain important information, which lead us to the better understanding of the relations between domestic leagues, international competitions and the players' age. Furthermore, it is crucial to verify the randomness of our sample by comparing the dataset's composition to what was empirically expected.

Thus, we may observe the following:

- the number of goalkeepers is much lower than the number of outfield players
- the age category including more observations is "24-26", while moving to extreme values (towards 16 or towards 31+), the number of observations decreases
- the number of the players included in the dataset are almost equal among the domestic leagues
- the minutes played increase when a player is getting closer to 30 years old, while when joining the 31+ age category, minutes tend to decrease
- similarly, the player's performance index is improving, as the player is getting older and plays more games, while when getting closer to his career termination his performance deteriorates
- the ratio between observations having participated in an international competition is just above 40%, which is realistic, as in most of the examined domestic leagues 30% of the teams qualify for next season's international competition

We, also, present the following graphs, which demonstrate the relation between the players' market value, and the way it has changed during the seasons of interest. Additionally, we examine the possibility of an increasing market value trend, by creating a graph, which illustrates the per cent alterations during each season.

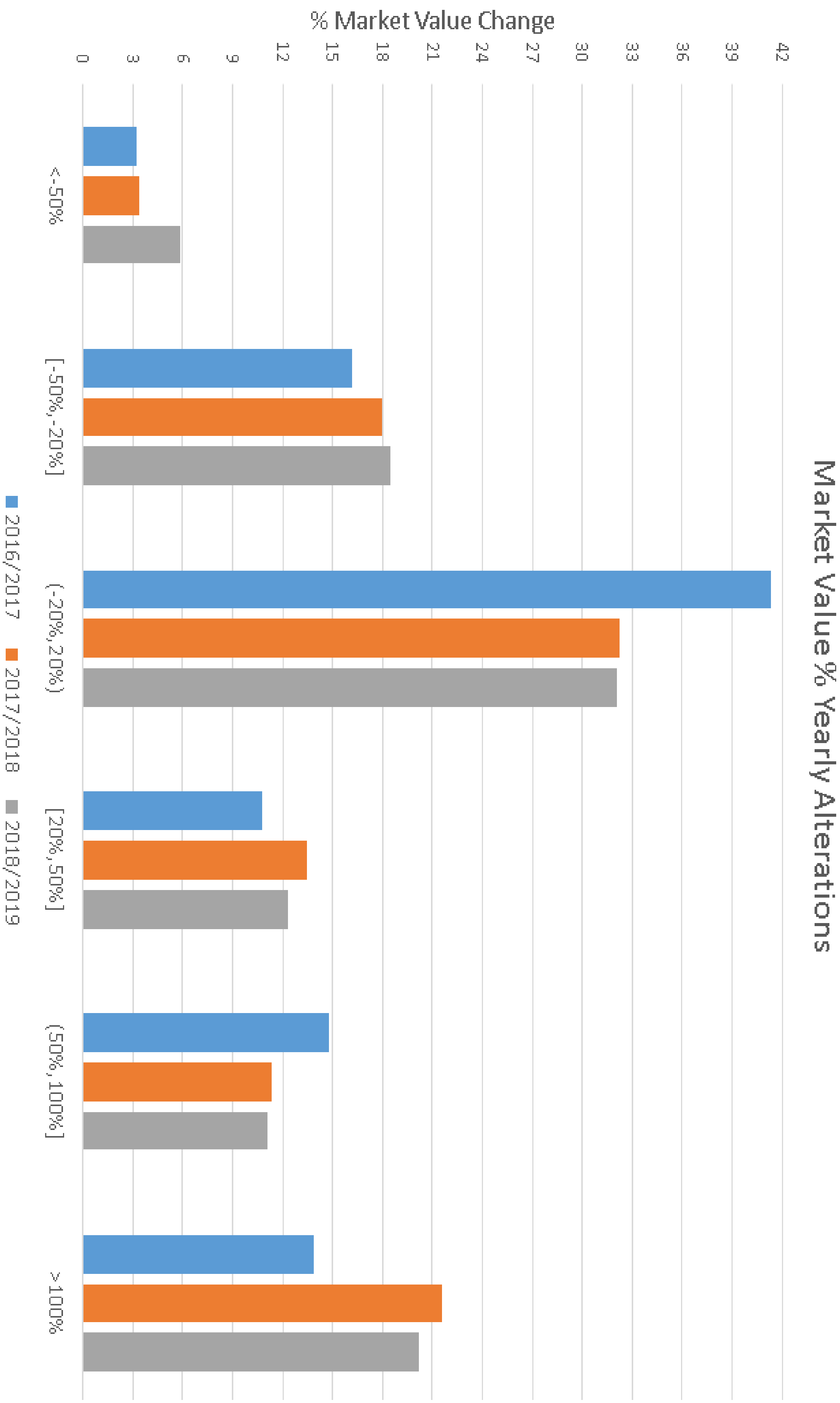

Figure 3.2: Market Value % Alterations by Season

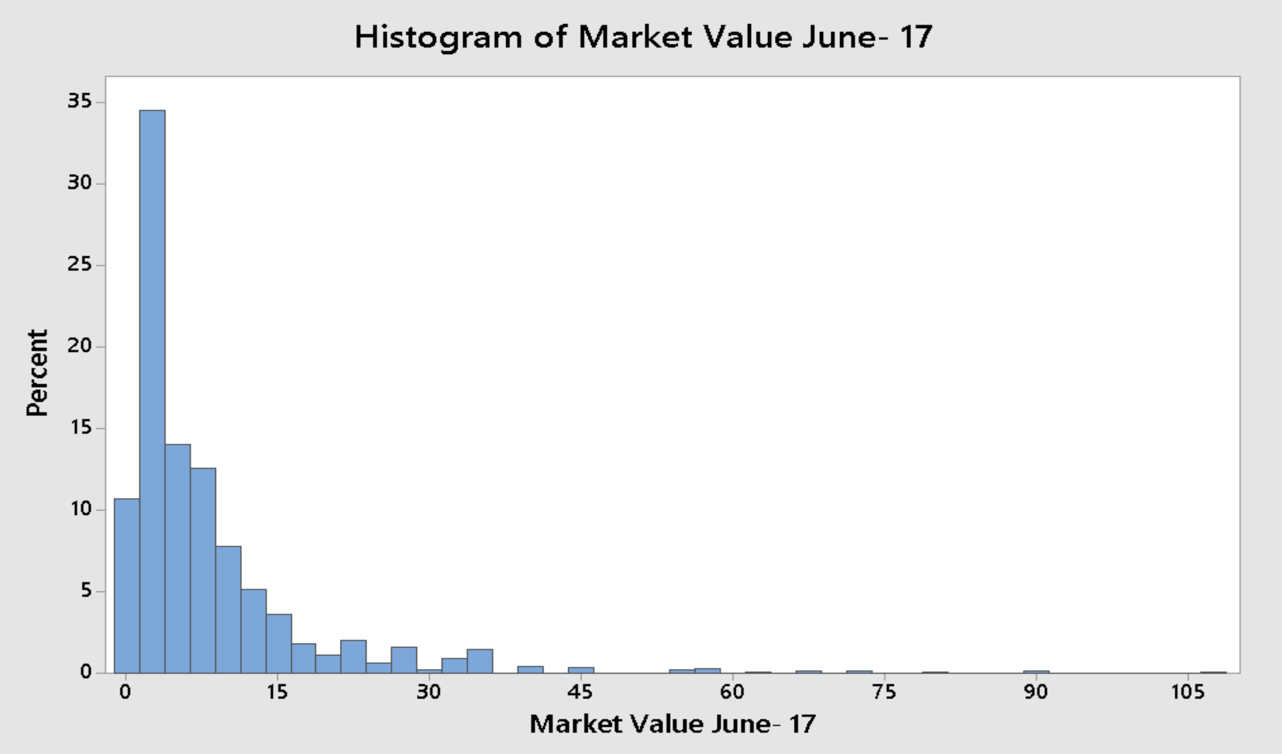

Figure 3.3: Histogram: Market Value on June 2017

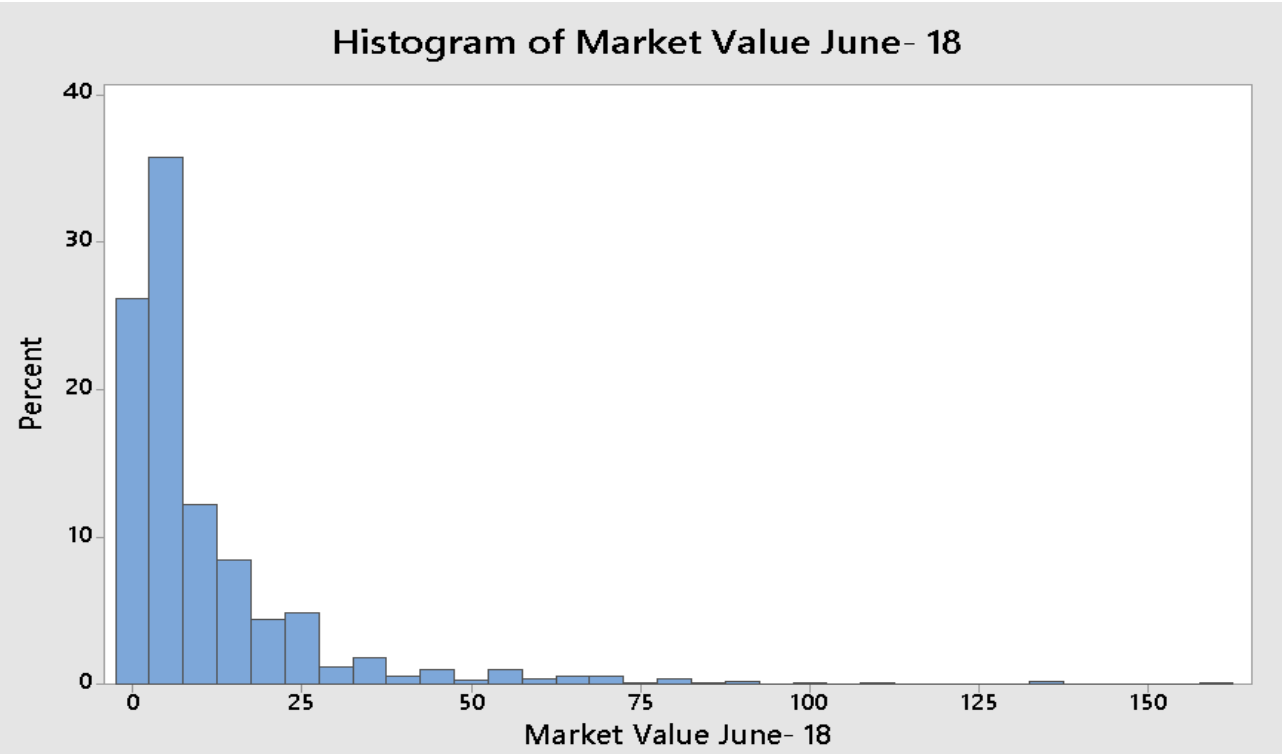

Figure 3.4: Histogram: Market Value on June 2018

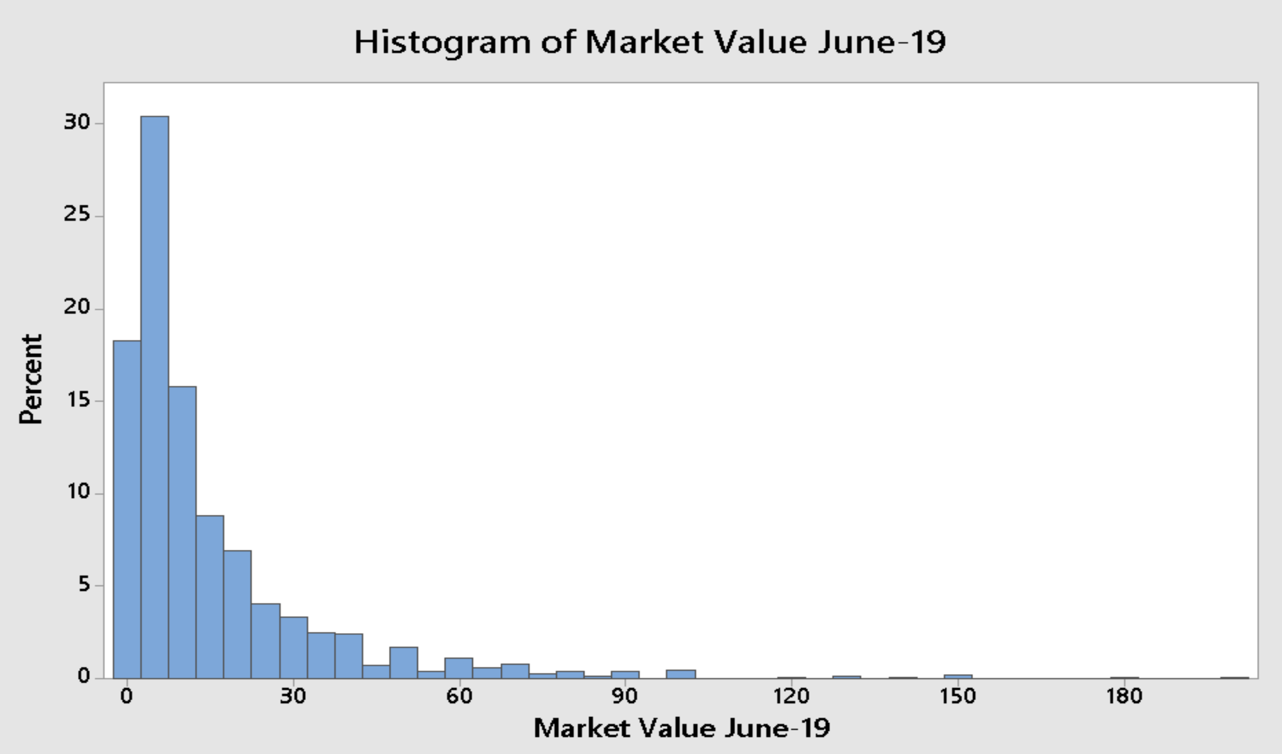

Figure 3.5: Histogram: Market Value on June 2019

From the graphs above, we observe that crowds influenced by the general transfer fee inflation tend to provide higher estimations regarding the players' market values. On figure 3.1, it is evident that, while during season 2016/2017 the main market value alterations were bounded between -20% and +20%, the next two seasons, we move on to higher alterations, thus there is a significant increase on the percentages for higher alterations. Additionally, from the histograms presented, it is clear that market values at the end of the season, tend to "stretch" towards higher values. From this observation, we conclude that players who used to perform close to the average values are led to higher values, while elite player keep on increasing their own market values.

On the following, we explain the reasons why this trend makes very complicated the analysis of market value alterations in football.

3.3 'Why is it so difficult?'

Since we have presented the dataset, we are almost ready to reveal MaVAM. However, at this point, it is important to explain a few reasons, why modelling a player's market value alteration is so complex.

3.3.1 Statistical Difficulties

These justifications derive from the initial tests performed in order to conclude to the final model. The main issue was to identify the appropriate way to express the relation between the model's response (Player's Market Value at the end of the year) and the features, which we empirically know they affect it. For example, the player's performance index is one of the major parameters that may alternate the player's market value either upwards or downwards. As a result, there can be no model that takes into account neither a specific performance index nor the stats (goals, assists, passes etc.) that are used to determine it. However, as the performance index is the mean rating that the player receives throughout the season, there is no difference between a player having played, for example, 3000 minutes, and someone who has played for only 90 (one game). Thus, we understand that every variable used in the model, must be defined in the proper way, so as to be considered statistically important, apart from the role we empirically know.

3.3.2 Data Inconsistencies

The second main reason that this process is difficult lies on the data available for research. As we have already explained, the assets (players) do not have a specific price tag (market value), so even though, our data source, transfermarkt.de, is a very efficient platform via its own crowd sourcing method, our data cannot be by nature absolutely exact, as the ones for stock prices. For example, we mention that Lionel Messi's Market Value during 2016-2017 remained the same (at approximately 120 million euros), while during 2017-2018, it increased by 50 %, even though his football performance remained approximately unchanged.

In addition, as already mentioned, we observe that there is a tension to ”stretch” the range between the market values of the players.

It is evident that because of the football evolution itself (injury rehabilitation, increased appearances from young players, sponsorship deals), but also due to the price inflation, the difference between the average players and the elite ones is growing rapidly.

Finally, we underline that MaVAM uses only performance related features, so it effectively estimates the market value alterations that occur, due to the player's achievements on the pitch. Even though, we perceive that a player's popularity can change his market value, especially fan wise, it is not easy to detect the popularity data that origin from the player's performance. However, in this work, our goal is to estimate the performance caused alteration, which will be beneficial for the club, but without being discouraging towards a marketing perspective future add-on.

3.4 Final Preparations

In this section, we present the pre-processing steps, we had to use on our data to prepare them for their addition to our model.

As already mentioned, an important element of this work is to distinguish leagues and clubs by including the UEFA Country and Club Coefficients. Firstly, we assume that the league's effect is stronger than the club's, and as a result, we obtain:

$$TotalCoefficient = \frac{UEFAClubCoefficient + 3 * UEFACountryCoefficient}{4}$$

For example, the Total Coefficient for the football players who played during season 2016/2017 for Borussia Dortmund in German Bundesliga is derived from the UEFA Coefficient of Borussia Dortmund and UEFA Coefficient of German Bundesliga before the new season starts. We should, also, note that since our dataset consists of leagues and clubs that have zero UEFA Coefficient (either second tier leagues and clubs or South American), we have to handle them as following:

- Second Tier: We create a ratio for the average player market value $\frac{SecondTierAverageMarketValue}{TopTierAverageMarketValue}$, which is multiplied with the Country's UEFA Coefficient, and thus, we obtain the equivalent UEFA Coefficient for the second tier

- South American Leagues and Clubs: We use interpolation based on the average player market value between leagues that have UEFA Country Coefficient, and thus we obtain the equivalent UEFA Coefficient for the South American League. For reasons of simplicity, we assume that South American Clubs get the same Coefficient as the League they compete in.

The existence of clubs that have zero UEFA Club Coefficient, also, reinforces our assumption for Total Coefficient, as it stabilises the gap between the elite clubs and the weaker ones. Furthermore, while performing regression, this relation can be considered optimal, as it occurred via testing for weights from 0 to 1 for each UEFA Coefficient.

The second crucial aspect, we focus on, is the separation of International Competition (UEFA Champions League, UEFA Europa League) minutes played and performance. To calculate the Total Coefficient, we use the aforementioned formula, where UEFA Club Coefficient is already known, and International Competition Coefficient (equivalent to UEFA Country Coefficient) is set as the average of UEFA Club Coefficients of the clubs that participated each year in the group stage. The logic behind this assumption is to dynamically change every year's competition Coefficient based on the participating clubs. From a practical point of view, a competition including stronger clubs will be more noticeable, providing higher UEFA Coefficient, and therefore, having a greater impact to the players' market values.

3.5 "Hello, I am MaVAM!"

In this section, we present MaVAM, the model developed to estimate the player's market value alteration during a football season and, also, compare it, with two other less sophisticated models.

3.5.1 Candidate Models Specification

It is our aim to prove statistically that the inclusion of UEFA Coefficients improves the model's results. We use three candidate models, which are presented below:

- Neither Country nor Club Coefficient (Plain Model)
- Only Country Coefficient (League Model)
- Both Club and Country Coefficient (MaVAM)

The other features used in all the models are the following:

- Player's Age
- Player's Market Value at the start of the season
- Domestic league minutes and performance index
- UEFA Champions League minutes played and performance index
- UEFA Europa League minutes played and performance index

The above features are used as described:

- Player's Age ⇒ Ageˆ2
- Player's Market Value at the start of the season ⇒ $\log OldMarketValue$

- Domestic league minutes and performance index $\Rightarrow$

$$TotalDomesticPerformance = \frac{DomesticMinutesPlayed \times DomesticPerformanceIndex}{TotalDomesticMinutesAvailable}$$

- International minutes played and performance index$\Rightarrow$

$$TotalIntern.Performance = ChampionsLeaguePerformance + EuropaLeaguePerformance$$

where, each competition performance is calculated as:

$$CompetitionPerformance = \frac{MinutesPlayed \times PerformanceIndex}{TotalMinutesAvailable}$$

Finally, we note that in the models that take into consideration Leagues or Clubs, we multiply their Coefficient with each corresponding Total Competition Performance.

3.5.2 Plain Model

We, now, present the best fitted model that used the aforementioned features without the use of UEFA Coefficients. Thus, the plain model does not distinguish players in terms of league or club.

The obtained model via regression is as follows:

Term	Coef	SE Coef	T-Value	p-value
Constant	1,8497	0,0397	46,61	0,000
$\log OldMarketValue$	0,74229	0,00591	125,64	0,000
Age^2	-0,000416	0,000015	-28,33	0,000
$\frac{DomesticPerformance}{Age}$	33,124	0,834	39,72	0,000
$\frac{ChampionsLeaguePerformance}{Age}$	21,56	1,48	14,55	0,000
$\frac{EuropaLeaguePerformance}{Age}$	11,30	1,62	6,98	0,000

Table 3.5: Plain Model: Regression Coefficients

Additionally, the model summary, as already seen on the above tables:

S	R^2	R^2_{adj}	R^2_{pred}
0,205544	84,56%	84,54%	84,51%

Table 3.6: Plain Model: Model Summary

And, finally, the Regression Equation:

$$\log NewMarketValue = 1,8497 + 0,74229 \times \log OldMarketValue - 0,000416 \times Age^2 + 33,124 \times \frac{DomesticPerformance}{Age} + 21,56 \times \frac{ChampionsLeaguePerformance}{Age} + 11,30 \times \frac{EuropaLeaguePerformance}{Age}$$

Model Evaluation

We will know proceed to the final model evaluation in terms of variables statistical significance, collinearity and our empirical interpretation.

- Statistical

At first, we examine R^2_{adj} whose value is 84,54%. This value is more than satisfying and it verifies that the data are very well interpreted by our model. Additionally, R^2_{pred}'s value, which stands at 84,51%, confirms our model's predictive capability. As far as the dependent variables are concerned, we have to examine the p-values of each coefficient, which are produced by the T-test or F-test. We observe that for our variables, the p-value is less than 0,001, which indicates that they are all statistically significant and the null Hypothesis (H_o =0: b_i =0 ,i=1,2,3,4,5) can be rejected. [16]

- Multicollinearity and Correlation

After passing the statistical tests, we shall now proceed to examine if there are problematic relationships between the variables selected in our model. On Table 3.7, we present the multicollinearity test, using Variance Inflation Factors (VIF):

Term	Coef	VIF
Constant	1,8497	
$\log OldMarketValue$	0,74229	1,30
Age^2	-0,000416	1,19
$\frac{DomesticPerformance}{Age}$	33,124	1,25
$\frac{ChampionsLeaguePerformance}{Age}$	21,56	1,31
$\frac{EuropaLeaguePerformance}{Age}$	11,30	1,06

Table 3.7: Plain Model: VIF Values

Therefore, as far as multicollinearity is concerned, VIFi, i=1,2,3,4 is much lower than 10, which is the critical value, so our variables have no problems in terms of multicollinearity.

Additionally, we present the correlation matrix for the model's variables:

Term	Age^2	$\frac{DomesticPerformance}{Age}$	$\frac{ChampionsLeaguePerformance}{Age}$
$\frac{DomesticPerformance}{Age}$	-0,392		
$\frac{ChampionsLeaguePerformance}{Age}$	-0,059	0,197	
$\frac{EuropaLeaguePerformance}{Age}$	-0,100	0,140	-0,069

Table 3.8: Plain Model: Correlation Matrix

Since there is no correlation value greater than 0,4, there is not enough evidence to state a strong relationship between our variables. Not only that, but the majority of the correlation values tend to zero, which is known to be the ideal one.

- Empirical Interpretation

  As our dependent variable is measured on the logarithmic scale, the models' coefficients can be interpreted rouglhy as per cent changes. The coefficients of the log-transformed independent variables have to be interpreted as elasticities. For example, an 0,01 increase on $\frac{ChampionsLeaguePerformance}{Age}$ alters the Market Value by 33,124 %. Additionally, it is important to note that all the model's coefficients are positive, apart from Age^2, whose increase has a negative impact on Market Value. Finally, the coefficients among the performance show that Domestic League Performance has a greater impact on the player's market value than International Competitions, while Champions League Performance is almost twice as influential than Europa League. These results are also to be expected, as domestic league is a long term competition throughout the football season, while Champions League is more prestigious than Europa League.

3.5.3 League Model

We, now, present the best fitted model that used the aforementioned features, but using this time the UEFA Country Coefficients as multipliers for each competition performance. Thus, the league model distinguishes players in terms of league but not in terms of club.

The obtained model via regression is as follows:

Term	Coef	SE Coef	T-Value	p-value
Constant	2,3094	0,0395	58,44	0,000
$\log OldMarketValue$	0,68967	0,00605	114,00	0,000
Age^2	-0,000599	0,000013	-45,03	0,000
$\frac{UEFACoefficient \times DomesticPerformance}{Age}$	0,000020	0,000000	43,02	0,000
$\frac{UEFACoefficient \times ChampionsLeaguePerf.}{Age}$	0,000017	0,000001	18,32	0,000
$\frac{UEFACoefficient \times EuropaLeaguePerf.}{Age}$	0,000022	0,000003	8,45	0,000

Table 3.9: League Model: Regression Coefficients

Additionally, the model summary, as already seen on the above tables:

S	R^2	R^2_{adj}	R^2_{pred}
0,201510	85,16%	85,14%	85,12%

Table 3.10: League Model: Model Summary

And, finally, the Regression Equation:

$$\log NewMarketValue = 2,3094 + 0,68967 \times \log OldMarketValue - 0,000599 \times Age^2 + 0,000020 \times \frac{UEFACoefficient \times DomesticPerformance}{Age} + 0,000017 \times \frac{UEFACoefficient \times ChampionsLeaguePerf.}{Age} + 0,000022 \times \frac{UEFACoefficient \times EuropaLeaguePerf.}{Age}$$

Model Evaluation

We will know proceed to the final model evaluation in terms of variables statistical significance, collinearity and our empirical interpretation.

- Statistical

  At first, we examine R^2_{adj} whose value is 85,14%. This value is more than satisfying and it verifies that the data are very well interpreted by our model. Additionally, R^2_{pred}'s value, which stands at 85,12%, confirms our model's predictive capability. As far as the dependent variables are concerned, we have to examine the p-values of each coefficient, which are produced by the T-test or F-test. We observe that for our variables, the p-value is less than 0,001, which indicates that they are all statistically significant and the null Hypothesis (H_o =0: b_i =0 ,i=1,2,3,4,5) can be rejected. [16]

- Multicollinearity and Correlation

  After passing the statistical tests, we shall now proceed to examine if there are problematic relationships between the variables selected in our model. On Table 3.11, we present the multicollinearity test, using Variance Inflation Factors (VIF):

Term	Coef	VIF
Constant	2,3094	
$\log OldMarketValue$	0,68967	1,42
Age^2	-0,000599	1,02
$\frac{UEFACoefficient \times DomesticPerformance}{Age}$	0,000020	1,14
$\frac{UEFACoefficient \times ChampionsLeaguePerf.}{Age}$	0,000017	1,32
$\frac{UEFACoefficient \times EuropaLeaguePerf.}{Age}$	0,000022	1,06

Table 3.11: League Model: VIF Values

Therefore, as far as multicollinearity is concerned, VIFi, i=1,2,3,4 is much lower than 10, which is the critical value, so our variables have no problems in terms of multicollinearity.

Additionally, we present the correlation matrix for the model's variables:

Term	Age^2	$\frac{UEFACoeff.\times Dom.Perf.}{Age}$	$\frac{UEFACoeff.\times Ch.L.Perf.}{Age}$
$\frac{UEFACoeff.\times Dom.Perf.}{Age}$	-0,120		
$\frac{UEFACoeff.\times Ch.L.Perf.}{Age}$	0,002	0,203	
$\frac{UEFACoeff.\times Eu.L.Perf.}{Age}$	-0,039	0,104	-0,082

Table 3.12: League Model: Correlation Matrix

Since there is no correlation value greater than 0,3, there is not enough evidence to state a strong relationship between our variables. Not only that, but the majority of the correlation values tend to zero, which is known to be the ideal one.

- Empirical Interpretation

As our dependent variable is measured on the logarithmic scale, the models' coefficients can be interpreted roughly as per cent changes. The coefficients of the log-transformed independent variables have to be interpreted as elasticities. For example, an increase by one on $\frac{ChampionsLeaguePerformance}{Age}$ alters the Market Value by 0.002 %. This percentage may seem negligible, but since the performance on this model is multiplied by UEFA Country Coefficient, which for example, for German Bundesliga is approximately 70,000, we understand that for a 25 year old player who participates in this league, the actual per cent alteration for an increase by one in performance is: $\frac{70{,}000\times 0.002\%}{25} = 5.6\%$. Additionally, it is important to note that all the model's coefficients are positive, apart from Age^2, whose increase has a negative impact on Market Value. Finally, the coefficients among the performance are not clear at first sight, regarding which Competition affects Market

Value the most. However, we know that top-5 Leagues have higher coefficients than both International Competitions. The Europa League's Coefficient is higher than the others, as it is much lower than them (approximately 23,000). Thus, the slightly greater impact on market value is counterbalanced by the low UEFA Coefficient Equivalent.

3.5.4 MaVAM

We, now, present the best fitted model that used the aforementioned features, but using this time the Total Coefficients as multipliers for each competition performance. Thus, MaVAM distinguishes players both in terms of league and club.

The obtained model via regression is as follows:

Term	Coef	SE Coef	T-Value	p-value
Constant	1,9675	0,0362	54,32	0,000
$\log OldMarketValue$	0,72807	0,00559	130,15	0,000
Age^2	-0,000441	0,000014	-31,91	0,000
$\frac{TotalCoefficient \times DomesticPerformance}{Age^2 \times OldMarketValue}$	0,000575	0,000012	46,15	0,000
$\frac{TotalCoefficient \times ChampionsLeaguePerf.}{Age^2 \times OldMarketValue}$	0,000781	0,000068	11,52	0,000
$\frac{TotalCoefficient \times EuropaLeaguePerf.}{Age^2 \times OldMarketValue}$	0,000901	0,000086	10,49	0,000

Table 3.13: MaVAM: Regression Coefficients

Additionally, the model summary:

S	R^2	R^2_{adj}	R^2_{pred}
0,198889	85,55%	85,53%	85,46%

Table 3.14: MaVAM: Model Summary

And, finally, the Regression Equation:

$$\log NewMarketValue = 1,9675 + 0,72807 \times \log OldMarketValue - 0,000441 \times Age^2 + \\ + 0,000575 \times \frac{TotalCoefficient \times DomesticPerformance}{Age^2 \times OldMarketValue} + 0,000781 \times \frac{TotalCoefficient \times ChampionsLeaguePerf.}{Age^2 \times OldMarketValue} + \\ 0,000901 \times \frac{TotalCoefficient \times EuropaLeaguePerf.}{Age^2 \times OldMarketValue}$$

Model Evaluation

We will know proceed to the final model evaluation in terms of variables statistical significance, collinearity and our empirical interpretation.

- Statistical

  At first, we examine R^2_{adj} whose value is 85,53%. This value is more than satisfying and it verifies that the data are very well interpreted by our model. Additionally, R^2_{pred}'s value, which stands at 85,46%, confirms our model's predictive capability. As far as the dependent variables are concerned, we have to examine the p-values of each coefficient, which are produced by the T-test or F-test. We observe that for our variables, the p-value is less than 0,001, which indicates that they are all statistically significant and the null Hypothesis (H_o =0: b_i =0 ,i=1,2,3,4,5) can be rejected. [16]

- Multicollinearity and Correlation

  After passing the statistical tests, we shall now proceed to examine if there are problematic relationships between the variables selected in our model. On Table 3.15, we present the multicollinearity test, using Variance Inflation Factors (VIF):

Term	Coef	VIF
Constant	1,9675	
$\log OldMarketValue$	0,72807	1,24
Age^2	-0,000441	1,12
$\frac{TotalCoefficient \times DomesticPerformance}{Age^2 \times OldMarketValue}$	0,000575	1,34
$\frac{TotalCoefficient \times ChampionsLeaguePerf.}{Age^2 \times OldMarketValue}$	0,000781	1,01
$\frac{TotalCoefficient \times EuropaLeaguePerf.}{Age^2 \times OldMarketValue}$	0,000901	1,03

Table 3.15: MaVAM: VIF Values

Therefore, as far as multicollinearity is concerned, VIFi, i=1,2,3,4 is much lower than 10, which is the critical value, so our variables have no problems in terms of multicollinearity.

Additionally, we present the correlation matrix for the model's variables:

Term	Age^2	$\frac{TotalCoeff. \times DomesticPerf.}{Age^2 \times OldMarketValue}$	$\frac{TotalCoeff. \times Ch.L.Perf.}{Age^2 \times OldMarketValue}$
$\frac{TotalCoeff. \times DomesticPerf.}{Age^2 \times OldMarketValue}$	-0,308		
$\frac{TotalCoeff. \times Ch.L.Perf.}{Age^2 \times OldMarketValue}$	-0,056	0,074	
$\frac{TotalCoeff. \times Eu.L.Perf.}{Age^2 \times OldMarketValue}$	-0,063	0,023	0,009

Table 3.16: MaVAM: Correlation Matrix

Since there is no correlation value greater than 0,4, there is not enough evidence to state a strong relationship between our variables. Not only that, but the majority of the correlation values tend to zero, which is known to be the ideal one.

- Empirical Interpretation

As our dependent variable is measured on the logarithmic scale, the models' coefficients can be interpreted roughly as per cent changes. The coefficients of the log-transformed independent variables have to be interpreted as elasticities. For example, an increase by one on $\frac{ChampionsLeaguePerformance}{Age}$ alters the Market Value by 0.0781 %. This percentage may seem negligible, but since the performance on this model is multiplied by Total Coefficient, which for example, for a Real Madrid player is approximately 150,000, we understand that for a 25 year old Real Madrid player, with initial market value 10,000,000 euros, the actual per cent alteration for an increase by one in performance is: $\frac{150{,}000 \times 0.0781\%}{625*10} = 1.871\%$. Additionally, it is important to note that all the model's coefficients are positive, apart from Age^2, whose increase has a negative impact on Market Value. Finally, the coefficients among the performance are not clear at first sight, regarding which Competition affects Market Value the most. However, we know that top-5 Leagues have higher coefficients than both International Competitions, and therefore, the slightly greater impact on market value is counterbalanced by their low UEFA Coefficient Equivalent.

3.6 Model Comparison

We shall, now, proceed to the comparison of the models created on the previous paragraphs. Examining the regression equations of all three models, we observe that the dependent variables related to performance are divided by Age, which complies with the intuitive assumption that older players are more unlikely to dramatically alter their market value than youngsters. Additionally, on MaVAM, performance factors are, also, divided by the initial Market Value, meaning that players with higher market values are obliged to perform in even higher standards to firstly maintain their current Market Value, and then, increase it.

On Table 3.17, we examine each model's S, R^2_{adj}, as well as, the in sample mean absolute error and mean absolute percentage error.

Model	S	R^2_{adj}	MAPE	MAE
Plain Model	0.205544	84.54 %	2.267074 %	0.151346
League Model	0.201510	85.16 %	2.2857012 %	0.152646
MaVAM	0.198889	85.53 %	2.2287017 %	0.149068

Table 3.17: Models' Statistical Comparison

Reviewing the models statistical comparison, as presented on Table 3.17, we conclude that MaVAM is the best model due to:

- its higher R^2_{adj}
- its lower S
- its lower MAE and MAPE Values

compared to the other models. Thus, we understand that not only MaVAM provides improved and detailed information regarding the player's future evolution in the market, but, also, it is superior statistically.

Finally, on the following tables, we present both MAE and MAPE for all three models, categorised by Age and League:

MAE by Age					
Model \ Age	16-20	21-23	24-26	27-30	31+
Plain Model	0.2224	0.1665	0.1411	0.1199	0.1295
League Model	0.2333	0.1672	0.1404	0.1198	0.1286
MaVAM	0.2145	0.1651	0.1400	0.1175	0.1295

Table 3.18: MAE by Age

MAPE by Age					
Model \ Age	16-20	21-23	24-26	27-30	31+
Plain Model	3.302	2.470	2.096	1.784	2.111
League Model	3.459	2.481	2.088	1.781	2.097
MaVAM	3.180	2.440	2.075	1.747	2.117

Table 3.19: MAPE by Age

The highlighted cells on the above table represent the lowest MAE and MAPE values, thus, the best fitted model for each age category. Firstly, we observe that MaVAM is the best model for 4 out of 5 age categories. The exception is for players aged 31+, where League Model provides slightly better results. Additionally, for players aged from 24 to 30, MaVAM's MAPE is close to 2 %, while for younger players MAPE increases with greater value 3.180% on 16-20 age category. However, this result was expected, as younger players' market values are characterised by great volatility and uncertainty.

MAE by League							
League / Model	EPL	SLL	GBD	ISA	FLU	2nd	RoW
Plain Model	0.1169	0.1826	0.1595	0.1440	0.1598	0.1453	0.1620
League Model	0.1156	0.1700	0.1538	0.1443	0.1625	0.1694	0.1870
MaVAM	0.1148	0.1738	0.1591	0.1449	0.1585	0.1439	0.1560

Table 3.20: MAE by League

MAPE by League							
League / Model	EPL	SLL	GBD	ISA	FLU	2nd	RoW
Plain Model	1.673	2.739	2.392	2.137	2.418	2.326	2.473
League Model	1.649	2.523	2.296	2.134	2.468	2.730	2.876
MaVAM	1.635	2.606	2.375	2.142	2.402	2.293	2.445

Table 3.21: MAPE by League

The highlighted cells on the above table represent the lowest MAE and MAPE values, thus, the best fitted model for each league. According to MAPE values, MaVAM is the best model for 4 out of 7 leagues, while League model is better for the rest. This implies that, even though, our UEFA Country Coefficient approach was initially correct, we have to improve in terms of the Club Coefficients. Additionally, the plain model's superiority in Italian Serie A shows that this particular league's coefficients should be revisited and improved. Regarding MAPE values for Spanish La Liga and German Bundesliga, which are relatively high compared to the other leagues, it probably occurs due to their corresponding UEFA Country Coefficients. To be exact, because of the Spanish domination in International Competitions during the past five seasons, the difference between Spanish La Liga and other leagues has stretched, without illustrating the real margin among them (numerical values available on Appendix). This is one of the most important reasons why UEFA Coefficients, even though are reliable, official information, should be balanced with other parameters, such as the League's and Club's Total Market Value. However, the better results compared to the plain model are evidence that their inclusion improves the effectiveness of the models.

3.7 Comments and Remarks

In this chapter, we examined the creation of a model that estimates the alteration of a player's market value during one football season, based on his performance in play. To sum up, the most important results are:

- MaVAM is more effective than models that do not take into consideration either the club that the player participated (League Model) or both the league and the club (Plain Model)

- UEFA Coefficients are efficient measures that offer useful conclusions but should, also, be revised and improved by the use of additional factors

Thus, we consider MaVAM as the best fitted model, when trying to estimate a yeraly market value alteration.

Chapter 4

Transfer fee Transformer Model

In this section, we present the second model developed for this work, which transforms the player's estimated next season's market value to the possible transfer fee.

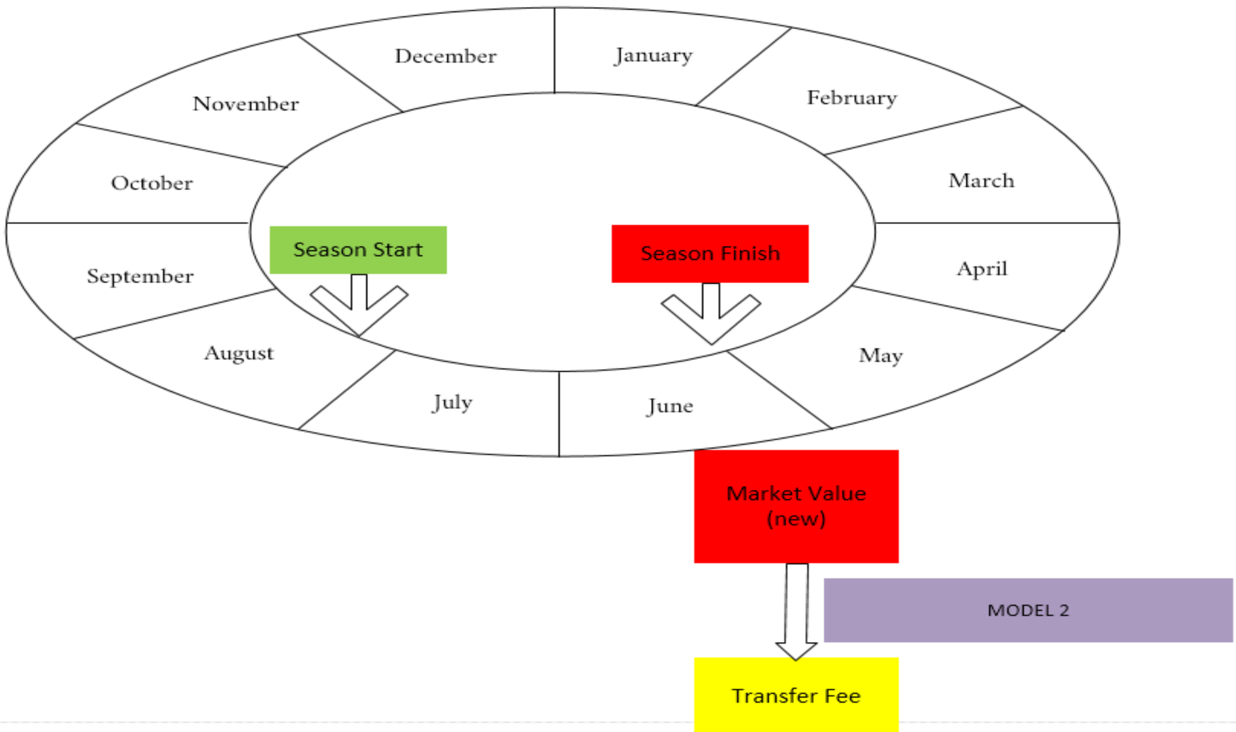

Figure 4.1: The Transfer Fee Transformer Model

4.1 The Dataset Used

We shall now proceed to the explanation of the dataset used for the model production. All the data were extracted from "transfermarkt.de", by reviewing each participating in TOP-5 European League player's transfer history. Transfers for which some of the required information was missing, were excluded from our study.

The data gathered for 1623 observations (completed transfers during the past 9 seasons) were the following:

- **Season** : The season where the transfer took place. It is noted that, for example, completed transfers on summer 2018 are registered on season 2018/2019.

- **Date** : The exact date when the transfer was completed. This information will be used in order to calculate the remaining days until the closure of the corresponding transfer window.

- **Position on the pitch** : The player's main position on the pitch. This factor's value will be represented in the models by dummy variables:

  - Goalkeepers: 0
  - Defenders: 1
  - Midfielders: 2
  - Attackers: 3

- **League Left** : The domestic league that the player's club participates

- **Club Left** : The name of the club that sold the player.

- **Club Left Total Market Value** : The selling club total market value at the time the transfer was completed in millions of €).

- **League Joined** : The domestic league that the player's new club participates

- **Club Joined** : The name of the club that bought the player.

- **Club Joined Total Market Value** : The buying club total market value at the time the transfer was completed (in millions of €).

- **Player's Market Value** : The player's market value at the time the transfer was completed (in €).

- **Transfer Fee** : The amount the selling club received as compensation for the transfer (in €).

- **Remaining Contract** : The contract's exact remaining time to expiration date. It is noted, that, these values were accurate even to the day, and expressed in years. For example, a remaining contract with value "1 year, 11 months and 25 days" was transformed to 1,9861.

- **Player's Age** : The exact player's age at the time the transfer was completed. It is noted, that, these values were accurate even to the day, and expressed in years. For example, a player aged "22 years, 7 months, 2 days" was transformed to 22,6638.

- **Days to Transfer Window Shutdown** : The days remaining until the transfer window closing.

- **Winter** : This value will be 0, if the transfer was completed during the summer transfer window, while 1, if the transfer was completed during the winter transfer window.

The total composition of the dataset is presented on the following tables and graphs:

Transfers' Categorisation by Age						
Season	16-20	21-23	24-26	27-30	31+	Total
18/19	49	128	134	102	27	440
17/18	44	98	118	82	15	357
16/17	36	90	92	68	12	298
15/16	33	73	90	41	4	241
14/15	12	39	38	20	0	109
13/14	16	27	22	6	2	73
12/13	10	17	13	4	1	45
11/12	14	26	18	2	0	60
Total	214	498	525	325	61	1623

Table 4.1: Transfers' Categorisation by Age

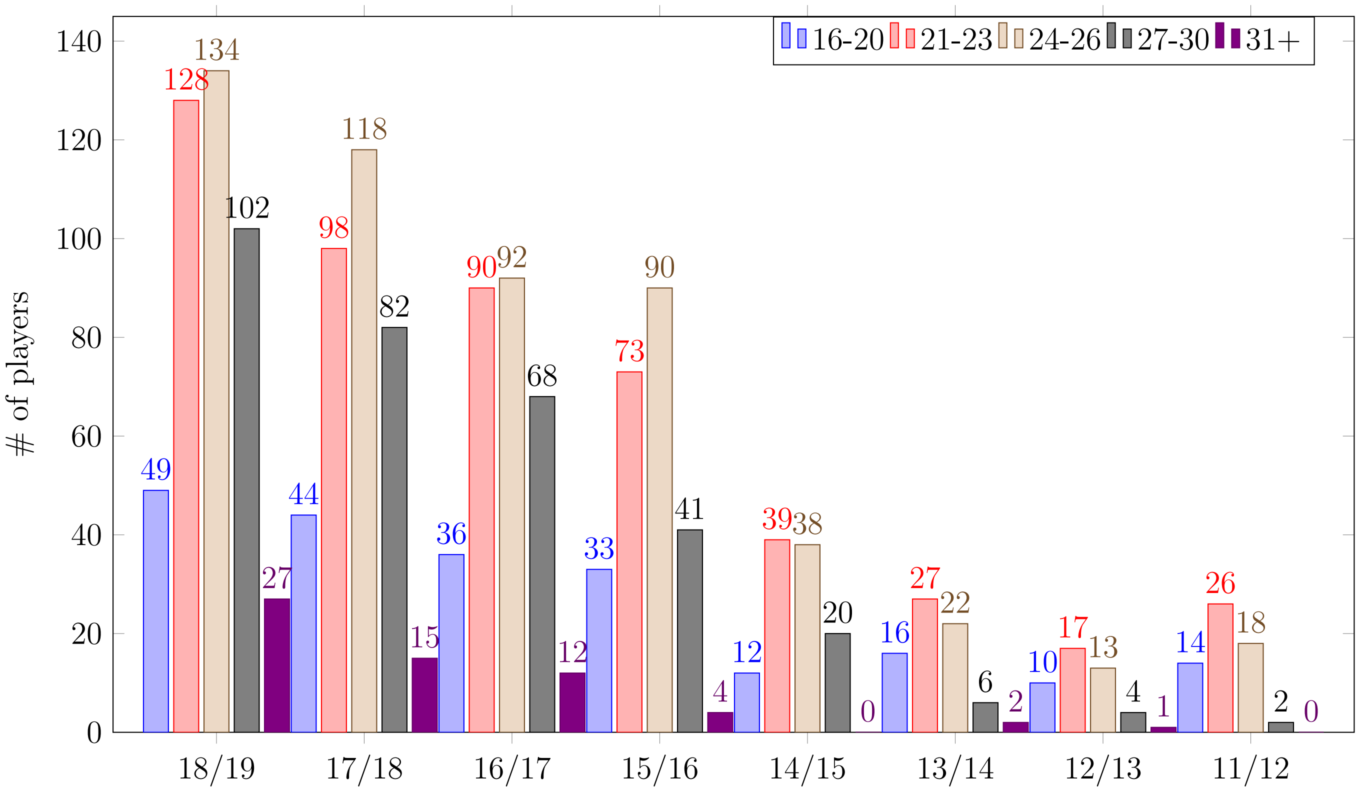

Figure 4.2: Transfers' Categorisation by Age

It is evident from the above table and figure that the majority of our data extracted, come from the past four seasons. Additionally, we observe that the biggest percent of the completed transfers occur when the player's age is in the range of 21 to 26 years. This fact was highly expected, as clubs invest mainly on young players due to their possible evolution in terms of potential and the probability of profit, in case they decide to sell them to another club. In this fashion, it does not come as a surprise, that only 61 from our 1623 observations were transfers of players whose age was 31 or higher. These players usually transfer as free agents, due to either the club's unwillingness to extend their contract or the players' desire to get a final profitable contract. On the other hand, players aged from 16 to 20 are more likely to be loaned to another club, so as to get more experience and playing minutes. Therefore, our observations on this age category are fewer compared to the others. Finally, the mean value of age for the average player on our dataset was 24,735 years.

Transfers' Categorisation by League							
Joined / Left	EPL	SLL	GBD	ISA	FLU	RoW	Total
EPL	98	18	13	18	16	65	228
SLL	42	52	10	46	17	18	185
GBD	27	11	86	9	10	11	154
ISA	29	14	12	108	7	14	184
FLU	40	15	18	13	60	12	158
RoW	140	80	127	102	82	183	714
Total	376	190	266	296	192	303	1623

Table 4.2: Transfers' Categorisation by League

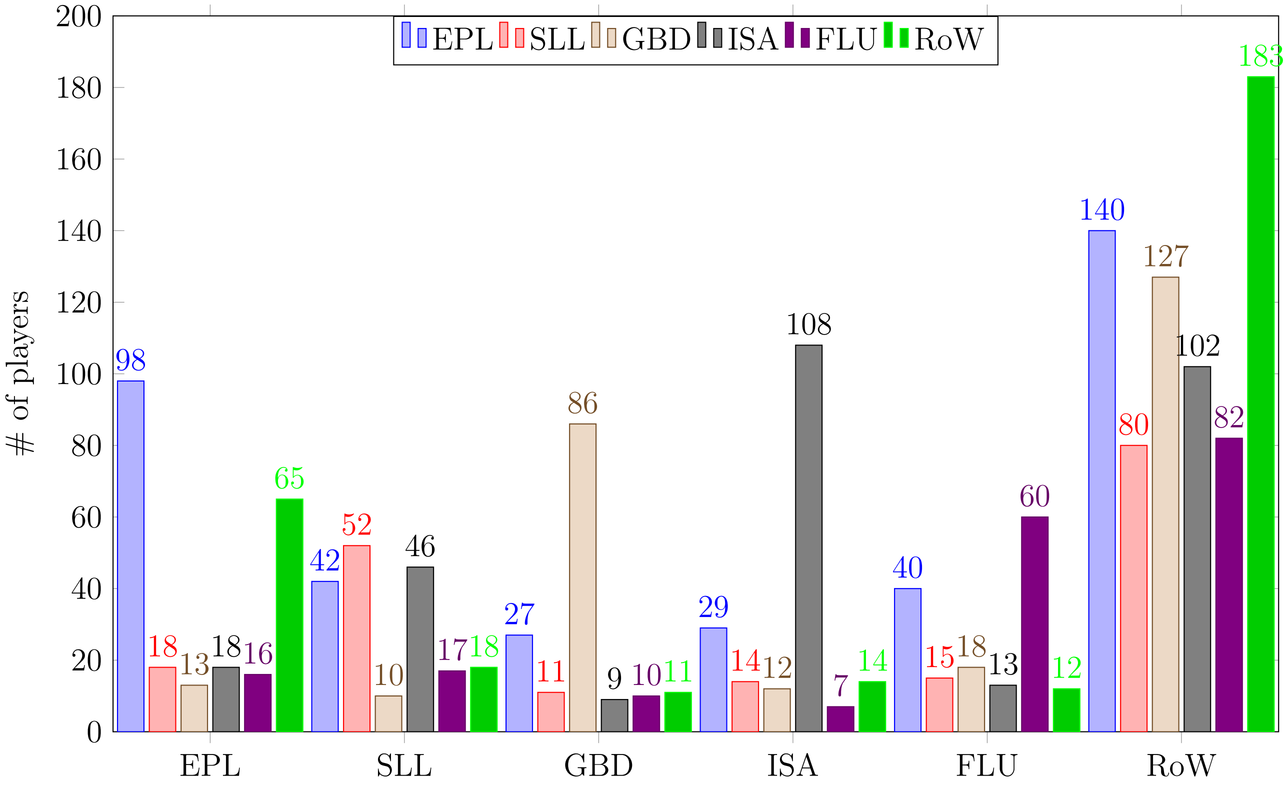

Figure 4.3: Transfers' Categorisation by League (x axis: League Left)

where:

- **EPL**: English Premier League
- **SLL**: Spanish La Liga
- **GBD**: German Bundesliga
- **ISA**: Italian Serie A
- **FLU**: French League Un
- **RoW**: Rest of World

From the figure and table for the transfers among the leagues, it is evident that clubs tend to sign players that used to play in the same domestic league. This can be explained by three main factors:

- It is easier for buying clubs to scout players that currently play in their own domestic league
- Buying clubs weaken their league's opponents by signing their best players
- Players do not need time to adjust to their new environment and the league's style of play

An impressive statistic element is that the English Premier League is by far the the domestic league that most of the players join. This mainly happens due to the English clubs capability to lure high profile players from the other leagues, by offering large contracts and occasionally paying great transfer fees to their corresponding clubs. On the other hand, regarding the league that the player left, there is a huge difference between the Rest of World departures and the ones from the Top-5 European Leagues. This was expected, as players from clubs participating in lower leagues in terms of prestige, tend to pursue a career to higher profile countries. However, this may create imbalances regarding the transfer fees paid to acquire players, as these dynamics have to be included in our model, as precisely as possible. For these relations, we use the Total Club Market Values, as they are presented on the following paragraph.

Transfers' Categorisation by Position						
Position	16-20	21-23	24-26	27-30	31+	Total
GK	4	16	37	35	12	104
Def	58	157	174	88	13	490
Mid	58	153	152	93	14	470
Att	94	172	162	109	22	559
Total	214	498	525	325	61	1623

Table 4.3: Transfers' Categorisation by Position

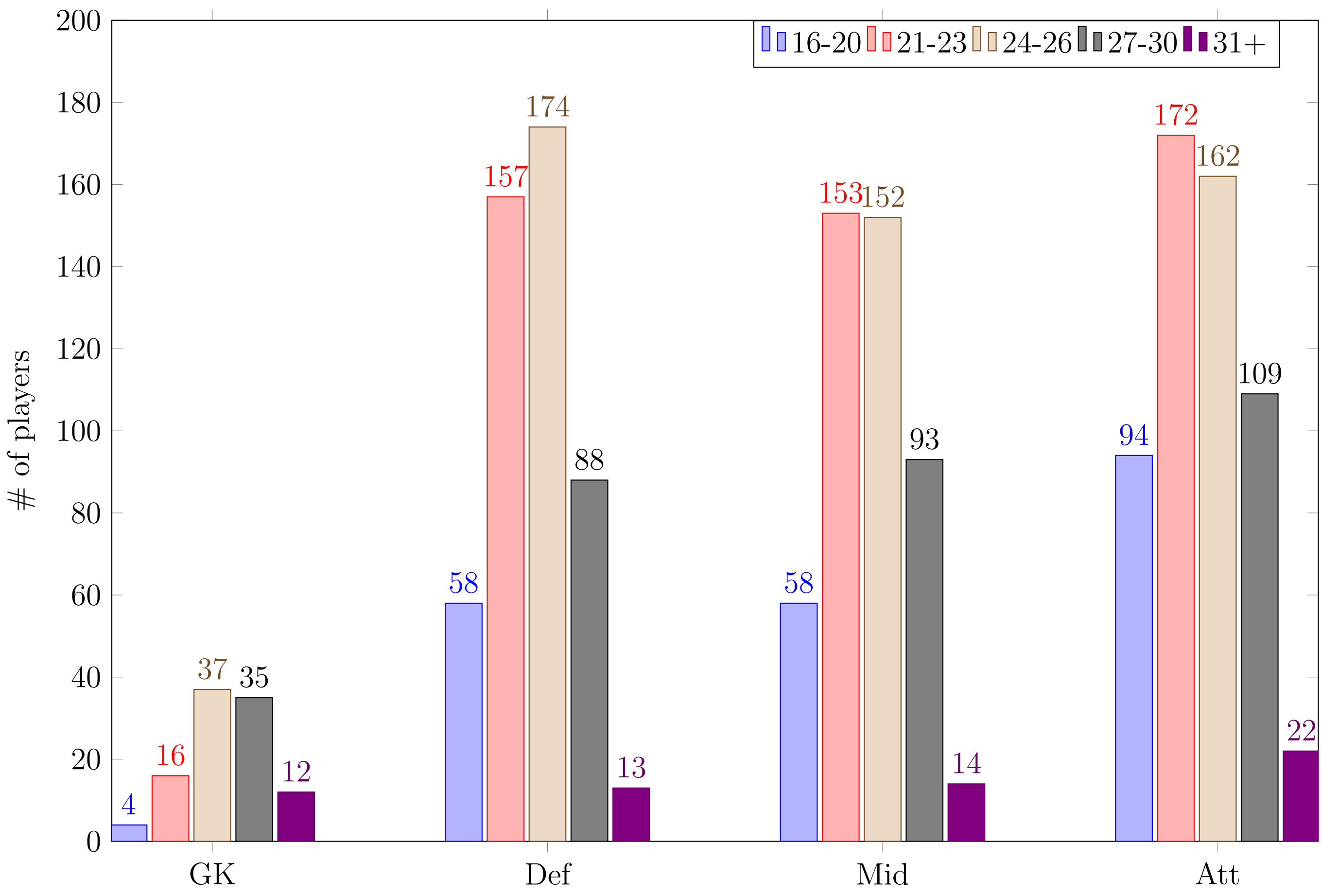

Figure 4.4: Transfers' Categorisation by Position

Finally, we present the composition of the dataset categorised by position on the pitch. We, firstly, observe that the number of goalkeepers transferred is lower compared to the other positions, which was expected as not only goalkeepers are fewer in terms of population, but, also, clubs are reluctant to sell a solid performance wise goalkeeper. On the other hand, players from other positions have similar observations in the dataset, and the main pool comes from players aged between 21 to 26 years. However, young attacking players, aged from 16 to 20, account for almost 6 % of the dataset, which shows the clubs' tendency to invest on evolving goalscorers.

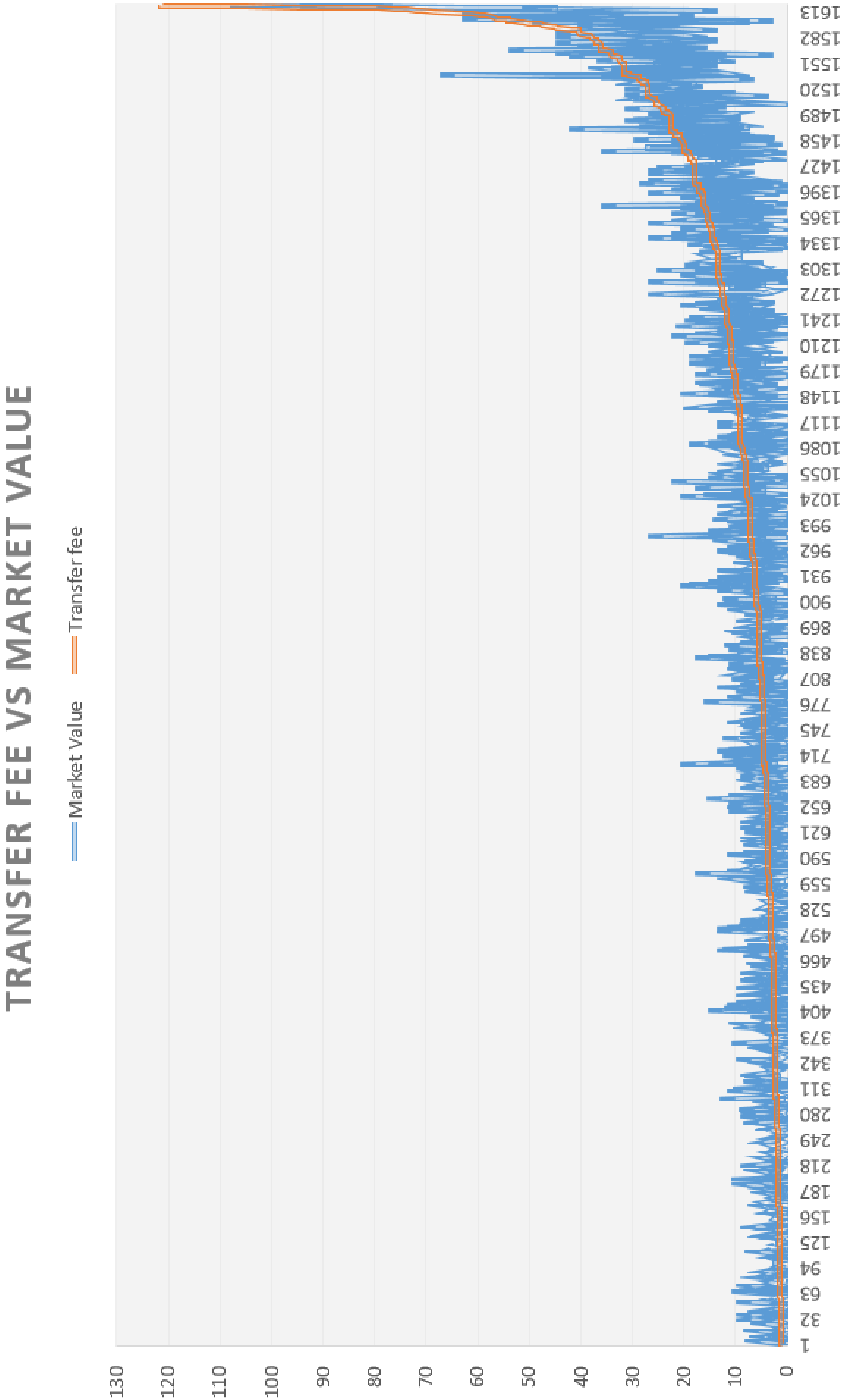

Figure 4.5: Transfer Fee vs Market Value

The final graph we present depicts the relation between the Transfer Fee, which was paid by the buying club and the Market Value of the player at the time the transfer was completed. It is obvious that there are many fluctuations around the corresponding transfer fees paid, which not only implies that there are many factors affecting their relation, but, also, that there is possibly some random elements originating from the clubs' decisions. On the next paragraph, we will try to identify these relations and create a model to transform a player's market value to a possible transfer fee.

4.2 Transfer Fee Transformer

We are, finally, ready to present the created model, which transforms the estimated market value to the possible transfer fee. To achieve the best result, we gradually add independent variables, while our response is $\log(TransferFee)$.

We begin by creating a baseline model (Model 1), where the dependent variable is only $\log(MarketValue)$. Model 2, also, includes the player's "Age" and "Remaining Years on the current Contract". Model 3 involves the dynamics between the negotiating clubs, where, as already mentioned, these relations are represented on our model with Total Club Market Values on the beginning of the summer. We should note, that, alterations occurring on this feature during a transfer window due to other deals, are not taken into consideration. This is an important aspect, as we these changes affect the club's financial status on the market (for example, via a very profitable deal), and consequently, makes our study even more complex. Model 4 includes the player's position, as well as the days that the transfer window remains open. FInally, we should not that "Winter" feature did not show any statistical significance, thus, it is omitted in our models.

4.2.1 The Model

On the following table, we present the obtained models via regression:

Response: log($TransferFee$)	Model 1	Model 2	Model 3	Model 4
Independent Variables				
Constant	2.2057***	2.6372***	3.3211***	3.3484***
	(0.0948)	(0.0931)	(0.0996)	(0.0985)
log($MarketValue$)	0.6863***	0.6851***	0.5509***	0.5508***
	(0.0143)	(0.0137)	(0.0158)	(0.0159)
Remaining Years		0.00613***	0.05834***	0.05727***
		(0.00613)	(0.00575)	(0.00568)
Age		-0.02243***	-0.01873***	-0.01859***
		(0.00194)	(0.00185)	(0.00185)
Total Club MV Left			0.000115***	0.000109**
			(0.000033)	(0.000032)
Total Club MV Joined			0.000551***	0.000569***
			(0.000039)	(0.000039)
Position				0.01448*
				(0.00611)
Days to Shutdown				-0.001526***
				(0.000252)
S	0.270166	0.248253	0.232969	0.229871
R^2	58.65 %	65.13%	69.33%	70.18%
R^2_{adj}	58.63%	65.07%	69.23%	70.05%
R^2_{pred}	58.53%	64.93%	69.08%	69.85%

Table 4.4: Regression Models

Notes: *p<0.05 , **p<0.01, ***p<0.001; standard errors in parentheses.

4.2.2 Models' Evaluation

We shall, now, proceed to the final model evaluation in terms of variables statistical significance, collinearity and our empirical interpretation.

Statistical

At first, we examine R^2_{adj}, whose value is 70,18%. This value is more than satisfying and it verifies that the data are very well interpreted by our model. Additionally, R^2_{pred}'s value, which stands at 69,85%, confirms our model's predictive capability. As far as the independent variables are concerned, we have to examine the p-values of each coefficient, which are produced by the T-test or F-test. We observe that for all our variables, p-value is less than 0,1, which indicates that they are all statistically significant and the null Hypothesis (H_o =0: b_i =0 ,i=1,2,3,4,5,6,7) can be rejected. [16]

Multicollinearity and Correlation

After passing the statistical tests, we shall now proceed to examine if there are problematic relationships between the variables selected in our model. On Table 4.5, we present the multicollinearity test, using Variance Inflation Factors (VIF):

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	3.3484	0.0985	34.01	0.000	
log($MarketValue$)	0.5508	0.0159	34.72	0.000	1.70
Remaining Years	0.05727	0.00568	10.08	0.000	1.10
Age	-0.01859	0.00185	-10.07	0.000	1.13
Total Club MV Left	0.000109	0.000032	3.35	0.001	1.15
Total Club MV Joined	0.000569	0.000039	14.67	0.000	1.50
Position	0.01448	0.00611	2.37	0.018	1.06
Days to Shutdown	-0.001526	0.000252	-6.05	0.000	1.02

Table 4.5: VIF Values

Therefore, as far as multicollinearity is concerned, VIFi, i=1,2,3,4,5,6,7 is way lower than 10, which is the critical value, so our variables have no problems in terms of multicollinearity.

Additionally, we present the correlation matrix for the model's variables:

Term	$\log(MV)$	Age	Left	Joined	Position	Remain. Years
Age	0.161					
Left	0.346	0.097				
Joined	0.544	-0.049	0.236			
Position	0.079	-0.103	0.002	-0.066		
Remain. Years	0.192	-0.182	0.063	0.167	0.013	
Days to Shutdown	0.063	-0.030	-0.016	0.050	-0.095	-0.017

Table 4.6: Correlation Matrix

Since there is no correlation value greater than 0,6, there is not enough evidence to state a strong relationship between our variables. Not only that, but the majority of the correlation values tend to zero, which is known to be the ideal one.

Empirical Interpretation

We will, now, examine our created model in terms of empirical evaluation, meaning that our results agree with the theoretical we expected. As already mentioned, on the equivalent paragraph on Chapter 3, using logarithms transforms the relations between our variables and the response to percentages.

To begin with, "Remaining Years" are multiplied with a positive coefficient, as when a player is bound by a contract, his club is highly likely to demand more money to release him. To be exact, a contract extension by one year increases the corresponding Transfer Fee by 5.72%. As far as "Age" is concerned, the negative sign was, of course, expected, due to the fact that, while players are getting older, they approach their career retirement. This observation verifies our statement on paragraph 4.1, which focused on the clubs' desire to invest more resources on young players. Additionally, when a player's age increases by one year, a decrease of 1.859 % on Transfer Fee should be expected. The coefficients for the club dynamics show that, while both tend to increase the value of the Transfer Fee, "Total Club Market Value Joined" tends to impact it in a greater way. Furthermore, "Position" is multiplied with a positive coefficient, meaning that attackers are more expensive than midfielders, while midfielders are more expensive than defenders. According to the coefficient, this increase is 1.448 %, for each step. Finally, "Days to Shutdown" is multiplied with a negative coefficient, which follows the market's tensions, as when the transfer window reaches its closure, selling clubs demand higher transfer fees, capitalising on the emergency situation the buying club has found itself. As a numerical example, we mention that a transfer completed at the start of the transfer period compared to one completed at the end may be cheaper up to 9.156%.

Thus, we conclude that even from an empirical point of view, our model corresponds satisfyingly with the theoretically expected results.

4.2.3 Unusual Observations: Residuals

In this section, we present some of the observations that our model did not effectively fit. Even though the average prercentage error for the whole dataset is 46.22%, we should note that as depicted on "Figure 4.5: Transfer Fee vs Market Value" there is an evident randomness between the two values, which is not easy to model, if possible at all. Furthermore, we should not neglect the fact that apart from the features we used in our model, decisions from clubs regarding transfers may be affected from other parameters such as immediate liquidity necessity, a player's imminent demand to depart or even the board's incompetency to negotiate the proper way. To support this allegation, we present five indicative examples, while, also, justifying why this transfer was not conducted in an effective way by either side.

Example 1			
Season	Age	Market Value	Remain. Years
18/19	25.5806	10.8m	0.833
Days to Shutdown	TMV Left	TMV Joined	Fee Paid
0	868.25m	841m	4.50m
Fitted Value	Actual Fit	Abs. Error	Rel. Error
7.3822	24.11m	19.61m	435.85%

Table 4.7: Unusual Observations: Example 1

- **Justification** : This defender's market value was 10.8m, when the transfer was completed, while the transfer fee paid was only 4.50m. Even though the player's contract was due to expire the next summer, there were elite clubs negotiating, meaning that a highest transfer fee could be demanded. A rational explanation may be that the selling club wanted to offload the player, thus, the transfer fee for this 25 year old player was much lower than the corresponding market value.

Example 2			
Season	Age	Market Value	Remain. Years
12/13	18.997	1.62m	5
Days to Shutdown	TMV Left	TMV Joined	Fee Paid
60	154.5m	34.6m	1.08m
Fitted Value	Actual Fit	Abs. Error	Rel. Error
6.69	4.88m	3.8m	352.13%

Table 4.8: Unusual Observations: Example 2

- **Justification** : This 19 year old attacker was sold on the summer of 2012 for a transfer fee close, but lower than his market value. He was bound by a long term contract, while the selling club was an elite one. A rational explanation may be that, he was not ready to perform for the club's first team, so he was sold in order to pursue his career in a less demanding club. However, he is currently valued at 9 millions, proving that his sale was a mistake at that time. Probably, loaning the player to another club should be a more appropriate solution.

Example 3			
Season	Age	Market Value	Remain. Years
17/18	18.7472	0.135m	1.417
Days to Shutdown	TMV Left	TMV Joined	Fee Paid
0	16.10m	57.25m	7.20m
Fitted Value	Actual Fit	Abs. Error	Rel. Error
5.9773	0.95m	-6.25m	86.81%

Table 4.9: Unusual Observations: Example 3

- **Justification** : This young attacking midfielder was sold for an extremely large amount, compared to his market value at the time. His contract's remaining duration can not justify this huge difference (more than 5,200 %), while the clubs' total market value was relatively low. This probably derives from bad negotiating strategy from the buying club. We should note that two years after the transfer, this young player is valued at only 4.5m.

Example 4			
Season	Age	Market Value	Remain. Years
18/19	32.8306	3.15m	0.869
Days to Shutdown	TMV Left	TMV Joined	Fee Paid
14	835m	94.13m	1.17m
Fitted Value	Actual Fit	Abs. Error	Rel. Error
6.5048	3.18m	2.01m	173,3%

Table 4.10: Unusual Observations: Example 4

- **Justification** : This experienced defender was sold below his market value, from an elite club to a second tier one. The selling club probably wanted to offload the player, as the manager considered that he was not needed in their roster. Thus, the club demanded a lower transfer fee to make the transfer's completion easier for the buying club.

Example 5			
Season	Age	Market Value	Remain. Years
18/19	25.0806	19.8m	3.833
Days to Shutdown	TMV Left	TMV Joined	Fee Paid
2	337.2m	974m	19.35m
Fitted Value	Actual Fit	Abs. Error	Rel. Error
7.7522	56.52m	37.17m	192,09%

Table 4.11: Unusual Observations: Example 5

- **Justification** : This final example derives from the case of a young attacker, who was transferred from a medium club to a top club. We observe that due to the dynamics of the negotiating clubs, but also, due to the player's remaining contract, our estimated value is much higher than the transfer fee actually paid. However, it is important to note, that, this deal would probably be completed with our estimated value, as there was a specific release clause on the player between the two clubs.

The above examples are only indicative, as there many similar cases included in the dataset. As already mentioned, modelling the relation between the market value and the transfer fee is particularly challenging, due to all the factors that are relative in every transfer deal. Even though, in this work, we included the most important empirical features that possibly influence the fees required for a transfer to be completed, we underline that each transfer may be affected by more specific parameters.

On the following table, we present the alteration of the average percentage error, when some of this outliers are excluded from our study:

Observ. Excluded	MAE	MAPE
0	3.33m	46.22 %
200	3.18m	31.86 %
400	2.85m	26.09 %
600	2.25m	21.28 %
800	1.61m	16.73 %

Table 4.12: Mean Absolute Error Alteration

Since there is not an equivalent model created throughout the available literature and taking into consideration all the aforementioned points throughout this chapter, we believe that our model is successful in both statistical and empirical testing and may be considered as a solid basis either for further improvements or for immediate application.

Chapter 5

Conclusion

In this final chapter, we present the overview of this work, including achievements, applications and future work that remains to be done.

5.1 Summary of Thesis Achievements

This thesis introduces a new financial instrument, based on the Derivatives Theory, to the football transfer market. To achieve that, we initially created a model, which connects the market value at the beginning of the football season to the corresponding at the end (Chapter 3). Furthermore, we developed a second model, which transforms the player's market value to a possible transfer fee with respect to the features explained on Chapter 4. Using these models, clubs have now the opportunity to estimate future transfer fees, having only to forecast the minutes that the player will play during the next season and his performance

It is not undue to state that this work improved many parts of the available literature, regarding players' market value estimation. Important factors verifying this allegation are the distinction among domestic leagues and clubs, and, also, the inclusion of international competition performance. In addition, the model we developed on Chapter 4 is introduced for the first time, as there is no evidence for its existence prior to this work.

Apart from the statistical point of view, the main goal of this thesis is to present this new financial product, which creates many opportunities for football clubs, as explained on Chapter 2.

5.2 Applications

There are two main categories of possible applications, deriving from this work. At first, it provides the clubs useful tools, which can be applied in order to estimate market value alterations and improve the quality of the investments in terms of better player evaluation. Apart from its main aim, this financial product can be created for investing in players' rights without actually acquiring the player. This fact could add features to the football transfer market comparable to the ones of a stock market. Finally, what would be of much interest is the creation of an online game, similar to "Fantasy Sports", where users may combine their passion for football, and, also their skills in finance.

5.3 Future Work

As mentioned at the end of the previous chapters, there are some improvements that still have to be done to achieve even better results. One important aspect is to improve our measures of domestic leagues and clubs dynamics, meaning that even though UEFA Country and Club Coefficients are satisfying measures, they have to be replaced by better ones, which will include the aforementioned coefficients. In addition, it would be interesting to produce methods that identify the players' popularity data deriving from the performance and add these features supplementary to our models. Furthermore, we should gather more data for periods up to three months, so as to create more short term models. This would enable the clubs to evaluate more frequently the player's financial value change.

The final and most important extension that needs to be implemented is the creation of the secondary market, where the created financial derivatives on football players can be traded.

It is of high importance that the clubs expand their portfolio including not only players or facilities, but, also, other financial investment products.

Bibliography

[1] Wikipedia, the free encyclopedia, en.wikipedia.org/wiki. [Online; accessed 28-May-2019].

[2] Statista, statista.com/statistics. URL `statista.com/statistics`. [Online; accessed 28-May-2019].

[3] FIFA. Fifa global transfer market report. 2019.

[4] Bbc, bbc.com/capital. URL `https://en.wikipedia.org/wiki/Transfer_market`. [Online; accessed 28-May-2019].

[5] Bloomberg. 2019.

[6] Uefa, financial fair play: all you need to know, "uefa.com". 2015.

[7] John C. Hull. *OPTIONS, FUTURES, AND OTHER DERIVATIVES*. 2012.

[8] Howard Viney Radu Tunaru, Ephraim Clark. An option pricing framework for valuation of football players. *Review of Financial Economics*, 2005.

[9] Henning Kreis Stephen Herm, Hans-Markus Callsen-Bracker. When the crowd evaluates soccer players' market values: Accuracy and evaluation attributes of an online community. *Sport Management Review*, 2014.

[10] Arno Knobbe Miao He, Ricardo Cachucho. Football player's performance and market value. 2015.

[11] Markus Weinmann Oliver Muller, Alexander Simons. Beyond crowd judgements: Data driven estimation of market value in association football. *European Journal of Operational Research*, 2017.

[12] Sports illustrated, si.com/soccer. URL `RB Leipzig's Season Finish Could Wind Up Increasing Liverpool's Fee for Naby Keita`. [Online; accessed 30-May-2019].

[13] transfermarkt, .transfermarkt.de. URL `https://www.transfermarkt.de`. [Online; accessed 30-May-2019].

[14] marca, marca.com/en/football. URL `https://www.marca.com/en/football/barcelona/2016/08/20/57b5e92846163f7c028b45b6.html`. [Online; accessed 30-May-2019].

[15] Fifa, fifa.com. URL `https://www.fifa.com/governance/news/y=2019/m=2/news=chelsea-fc-the-football-association-sanctioned-for-international-transfers-of-mi.html`. [Online; accessed 30-May-2019].

[16] P. Oikonomou C. Caroni. *Statistical Regression Models*. 2017.