BI Norwegian Business School - campus Oslo

GRA 19502

Master Thesis

Component of continuous assessment: Thesis Master of Science

Market Efficiency in the European Football Transfer Market

Navn:	Christoffer Kroken, Guleed Hashi

Start:	02.03.2017 09.00
Finish:	01.09.2017 12.00

Guleed Hashi & Christoffer Kroken

Programme:
MSc in Business, Major Finance

Date of submission:
01.09.2017

Table of Contents

1 Abstract

This dissertation studies market efficiency in the European transfer market for football. Data on player statistics and transfer fees is collected over the five most recent seasons, and covers Premier League, Bundesliga, La Liga, Serie A and Ligue 1. With three linear regression models, we attempt to answer in which degree the market is efficient. The player rating model is meant to capture on pitch performance, and in turn used as an explanatory variable in the final transfer fee model. Two transfer fee models are then estimated based upon player- and market characteristics. Our findings show that there exists a premium for South American players, concluding that the market is weakly inefficient due to the premium on South American players.

2 Introduction

Our thesis is founded upon the principles of the Efficient Market Hypothesis, and focuses on the European transfer market for football. Football is the most viewed sport in the world, and large amounts of capital are involved when clubs invest in new players. The football industry has experienced profound growth in revenues in recent years. Revenues in terms of broadcasting rights has been flourishing, alongside with substantial investments in clubs and players (P. J. Sloane, 2015). The eagerness to stay competitive, and reap financial rewards, is observable in transfer fee- and wage records within the industry. As the report from FIFA TMS shows, the summer transfer window comprised of 7,325 international transfers, with a global spending amounting to USD 3.72 billion (*Big 5 - Transfer Window Analysis Summer 2016*, 2016). The current market depicts a higher activity level than ever before, and is observable as player prices have surged. The abundance of detailed player information, such as performance data, transfer fees and wages, makes football an interesting market to investigate various economic theories.(Frick, 2007).

The publication of Moneyball raised questions regarding the use of player performance data for valuation and recruitment in sports. Michael Lewis´ book tells the story about the ability of Billy Bean and the baseball club Oakland Athletics. Their ability in sustaining a competitive advantage eight years in a row, despite low wage expenditures, was highly intriguing. This new rationale of player recruitment and valuation got other sports industries more engaged in sports analytics to improve sporting decisions. Today, we have technology capturing detailed performance data, allowing clubs to assess athletes with quantitative methods that were unachievable some years ago.

Through investigating the relationship between observed transfer fees, player characteristics and on-field performance we will attempt to determine if the transfer market is efficient. Most earlier studies have either created a performance rating, or analyzed transfer fees, apart from (Sæbø, 2016). Using the work of Sæbø (2016) as inspiration, we will try to answer the research question.
Our research question is the following: "*Is the European transfer market for football efficient, and what determine transfer fees in football*?",

In the attempt to answer the question, we will construct three regression based models:

1. A player rating based on performance data which will capture on-pitch performance (Fixed effects method).
2. A transfer fee model to determine which player- and market characteristics that are significant in determining prices.
3. An augmented transfer fee model including the player rating to test for market efficiency.

In *section 2,* we start by describing the characteristics, history and mechanisms of the European transfer market, with focus on the top tier divisions in the UK, Germany, Italy, Spain and France. Moving forward, section 2 continues to investigate and discuss existing literature on economics of sports in addition to classical financial theories. *Section 3-5* discusses the theory, data and methodology applied in the empirical investigation. *Section 6* provides results and analysis from the study. In *section 7,* we discuss key findings and conclude on the empirical paper, with suggestions for further research steps.

3 Background and literature

To deepen the readers understanding, and clarify the basis for this dissertation, section 2 will explain the context in which the investigation will pursue. The first part will focus on the peculiar mechanisms of the European transfer market, limited to the five most popular leagues in European football, while the literature section will review earlier studies` findings and thoughts.

3.1 Mechanics and" The Big 5"

To deepen the readers understanding, and clarify the basis for this dissertation, section 2 will explain the context in which the investigation will pursue. The first part will focus on the peculiar mechanisms of the European transfer market, limited to the five most popular leagues in European football, while the literature section will review earlier studies` findings and thoughts.

The main objective of the player transfer market is outlined by F. Carmichael and Thomas (1993, p. 1467):

The primary purpose . . . should be twofold: (i) to facilitate and organise the acquisition and exchange of players by clubs to enable the reconstitution of teams with the aim of increasing playing strengths and improving team performances; (ii) to facilitate the movement of players between clubs in their search for better opportunities, higher earnings or increased job satisfaction.

The labor market in European football is irregular compared to a normal labor market (P. Sloane, 1969; P. J. Sloane, 1971).The main distinction is that football players (workers) change workplaces by the transfer of monetary amounts to the employer holding their contract. Even though irregularities exist, other aspects of traditional labor are similar. Many organized team sports employ a transfer system where players are traded between clubs. North American team sports trade players through draft picks and swap deals, while European sports trade players for cash settlements.

The transfer market in European football is characterized a free market. Player movements are always possible given that contractual obligation between clubs and players are satisfied(P. J. Sloane, 1971). Throughout the years, this system has forgone several changes to improve functionality, transparency and maintain competitive balance. Each country has its own National football Association (NA), regarded as an authority overseeing that clubs comply with rules and regulations. FIFA, the governing body of international football sits at the top, followed by UEFA, the governing body for European football. National Associations sets local rules, but needs to comply with rules and regulations set by FIFA and UEFA. Restrictions and regulations are necessary to keep a healthy market. Due to interdependency between clubs, regulations are in place to keep the distribution of high quality players even(Peeters & Szymanski, 2014).

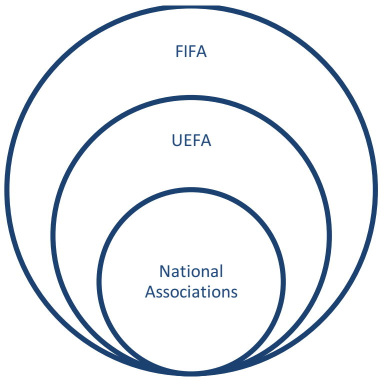


Figure 1: Hierarchical structure of governing bodies in football.

Players need registration in their respective National Association to play. To complete a transfer, players need approval from National associations of both the team they are joining and the team they are departing from. A major change in the transfer market is worth mentioning. The Bosman ruling of (1995) strongly improved the functionality of the transfer market. Before the ruling, free transfers was not possible without the consent of the club who held the contract, even if the contract was expired (Stephen Dobson & Goddard, 2001). This led to substantial bargaining power for clubs, restricting players' rights. The effect of the ruling entered force in 2001, enabling players who were in their last year of contract to move on a free transfer. After 2001, when the Bosman effect entered force, transfers were limited into two periods; summer- and winter window. This event was a pivotal point in the transfer market, shifting some of the bargaining power from clubs to players.

The construction of the leagues in big five in European football such as it encourages competition. Each league has twenty teams competing domestically against each other, except Bundesliga which have 18 teams. Teams compete for a position in the top three/four, depending on league, which qualifies for the UEFA Champions league.

3.2 Financial development of market

The commercialization of the sport has only increased with the years. With large prizes for winners of domestic- and European competitions, teams seek sporting- and financial success by acquiring top-level players. In 2016, the European transfer market recorded all-time high total spending. A total of 1,504 transfers between the top five European competitions and a total transfer value amounting to 2.748 billion dollars was recorded which was 74 % of total international spending (*Big 5 - Transfer Window Analysis Summer 2016*, 2016). Paul Pogbas world record signing in 2016, with a transfer fee of €105 mill to Manchester United, illustrates club willingness for success. Another interesting observation is that most transfers are without fees, consisting of loans and free transfers.

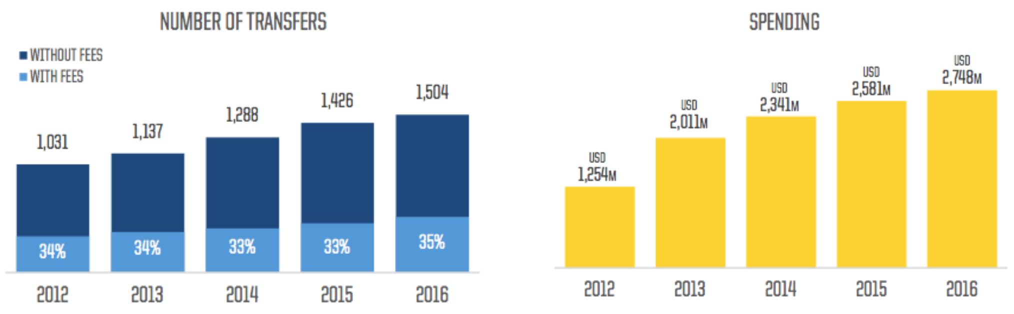


Figure 2: Number of transfers and total spending of the big five. *(Big 5 - Transfer Window Analysis Summer 2016, 2016)*

Financial strength is an important aspect for a clubs sporting success. The English Premier League has the highest revenue; followed by Germany, Spain, Italy and France. The major distinctions between the big five comes from broadcasting rights and sponsorships. This says something about the financial strength and ability to spend in the transfer market.

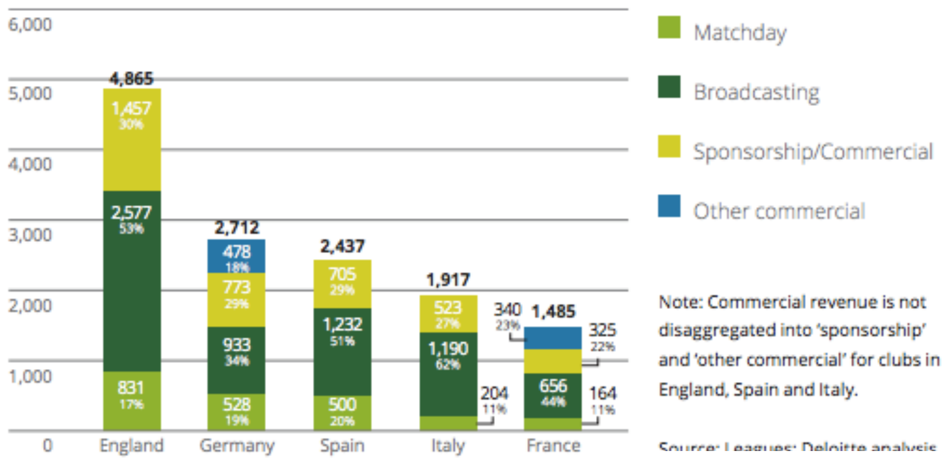


Figure 3: Total revenue by country. KILDE!!!!!!!

3.3 Literature review

The literature applied is comprised of classical economic papers in addition to papers on the economics of sports.

The literature and research published in relation to the economics of sports is nothing of news, as they can be dated all the way back to around the late 1960´s. Whilst football related studies were few in relation to other sports, studies on the subject has recently flourished. (Sloane, 2015). The scholars of studies conducted in the late 90´s seemed to agree on one thing; the determination of transfer fees could make better use of advanced player performance metrics (P. J. Sloane, 2015, p. 5).

Sloane was a pioneer in introducing the assumption that football clubs are utility maximizers, and not profit maximizers, a concept broadly recognized amongst his peers. Scholars (of the 1990´s) suggested that the field of study has been on standby until recent years (Fiona Carmichael, Forrest, & Simmons, 1999). Starting with Sloane´s study, regarding club behavior, we will present a literature review in relevance to our research area.

P. Sloane (1969) discusses the differences of North American and European team sports. To classify the primary business activity of football clubs, he analyzed the behavior of English football clubs. In doing so, he challenged Rottenberg ´s suggestion that football clubs are profit maximizers. Rottenberg saw no reason to treat professional sports leagues different from conventional business firms, hence the assumption of sports clubs being profit maximizers (P. Sloane, 1969). Sloane argues that that the assumption of profit maximization behavior is unfitting, because clubs operate with losses, while having limits on payment of fees to directors and dividends to shareholder. He concludes that clubs are utility maximizers with a financial constraint, and the utility function having two arguments; team performance and club profits.

S. Dobson and Gerrard (1999) developed a model estimating player transfer fees in the English premier league. The TP-CP model was derived from Sloane´s assumption. With the foundation of 1350 transfer fees between the period of June 1990 and August 1996, his model observes that the prices are mainly reflected by Player characteristics, selling-club characteristics, buying club characteristics and time effects. The result showed that these factors explain 79 % percent of the variations in real transfer fees. By also using the model to examine the inflation rate in the market, he concludes that statistically the market is highly rational.

Gerrard (2007) analyses the transferability of Billy Bean and Oakland Athletics success in sustaining a competitive advantage, over to the more dynamic and complex team sport of football. He identifies and discusses the issue with three main measurement problems in relation to copying the success of the Oakland A´s. This ultimately being the conceptual, technological and cultural barriers, concluding that the two first mentioned issues are difficult, but manageable. Further on, he emphasizes that the cultural barrier is the most enduring challenge in transferring sustainable competitive advantage to a dynamic sport such as football.

Leach and Szymanski (2015) reviews and challenges the generally adopted assumption that European football clubs are utility maximizers. Their study examines the performance of sixteen English football clubs that became listed during the mid-90s, and tracked the behavior of these clubs. Advocating that since the clubs were listed on a stock exchange, their behavior should drift towards a profit maximizing behavior. The study depicted no change in behavior after the listing, and they concluded that the result is more consistent with the assumption that clubs are more profit maximizing then first assumed.

More recent related papers are those of Frick (2007), Sæbø (2016) and McHale, Scarf, and Folker (2012). Frick´s (2007) paper reviews vast available evidence and models on the transfer market of football. He summarizes and compares different methodologies and results of other scholars within the field. McHale et al. (2012) creates a player performance index for EA sports. Their index attempts to rate players objectively with a single score using six sub-indices. The result showed that they were successful in creating such an index, which we will partly attempt to recreate. Sæbø (2016) evaluates market efficiency within the transfer market, and will be used for inspiration. They create a regression based rating to use as input for analyzing transfer fees.

Even with the increase of literature in recent years, scholars still haven’t found an answer to the puzzle in relation to the peculiarities of the economics of football. Still, the assumption of clubs being utility maximizers carry major influence in this field. As Sloane, and several other scholars suggest, the economics of sport has only now started to flourish (P. J. Sloane, 2015, p. 5).

4 Theory

The theories introduced in this section will form the foundation for our analysis. As there are differences between the transfer- and stock market, direct application of classical economic theory proves difficult.

4.1 Efficient Market Hypothesis

Fama (1970, p. 387) stated that "A market in which prices always "fully reflect" available information is called "efficient"". The theory states that in an efficient capital market, prices should fully reflect available information. The empirical work conducted can be divided into three categories; *strong-form, semi-strong and weak form.* The condition of weak form efficiency is that market prices reflect historical information. Secondly, semi-strong form test if the prices adjust in terms of new information. The condition is fulfilled if all publicly available information (and historical) is immediately reflected in market prices. Finally, *strong-form* tests if investors have monopolistic access to any information. Prices should reflect all information, including historical, public and private. To obtain this degree of market efficiency, everyone must possess the same specific information as industry experts and insiders.

This market is not frictionless, as information is gathered by firms, and sold to clubs for a fee, so the definition of what available information is in this context may be ambiguous. We define publicly available information as information that is obtainable for all. We will only consider information from on-field contributions, seasonality and player characteristics such as age and nationality.

4.2 Bargaining and Human Capital

Bidding and bargaining between parties occur in all markets with supply and demand where market prices are hard to pinpoint. Nash (1950) states that "A two-person bargaining situation involves two individuals who have the opportunity to collaborate for mutual benefit in more than one way". Explaining when n-number of parties are involved in a bidding situation, it ends up as a two-person bargaining. He continues to explain how the bargaining problem is idealized by assuming rationality among bidders. By this simplification, Nash creates a simplified economic environment which may not always be directly applicable.

F. Carmichael and Thomas (1993) attempted to solve bargaining in the transfer market by building on Nash´s models. Using the Nash Bargaining Solution, defined as the transfer fee f^* that maximizes utility for both players (both the selling club and buying club), they attempted to derive the relationship between the two. Their conclusion ended in that bargaining in the transfer market is not symmetric between buyers and sellers.

Nash studied a zero-sum game, a game where the gains in utility of player A is directly offset by the utility losses of player. Von Neumann (2004) extended the model further by suggesting a nonzero sum game. In the nonzero sum game the assumptions and restrictions of the zero-sum game is eased, giving way to cooperation and mutual benefits between the parties. The main separation of the two is that an optimal solution can always be found, and the combined gains/losses can differ from zero. An example of a non-zero sum game is the infamous Prisoner´s Dilemma.

5 Data

The data we work with in this master thesis is unbalanced panel data. To develop our models, we have collected team-/player specific metrics from whoscored.com, and transfer fee data from soccernews.com. Performance data collected is based on the top five leagues in European football. The horizon sets over a period of eight seasons (2009/2010-2016/2017) for each league, with 380 matches per season (306 in Bundesliga). Two matches were canceled; one in Serie A (2011-2012 season) and one in Ligue 1 (2016/2017), but we find these omissions to be insignificant for our conclusion. Data is collected for both home- and away matches.

A total of 14,606 football matches was collected, with twenty-one variables. To manage the data more efficiently and hopefully get relevant results for active players we choose to include only the five most recent seasons (2012/-2013-2016/2017). Variables were kept as before, but the number of games was reduced to 9129. The twenty-one variables obtained describes discrete player contributions in each match, such as key passes, shots, crosses and dribbles. Assigning the variables to every player, we obtain a dataset of 343,700 observations for the player specific dataset.

Transfer fees contained 6248 transfer with the same time span as data collected for matches. The data contains transfers between (and within) the big five, and inbound transfer from other leagues. Transfers based on loan agreements or free agents are omitted as there is no monetary value collected. Transfers from the big five to other leagues are also omitted for consistency. Players without game time will also be omitted. After this reduction, the dataset contained 1185 transfers.

The unbalanced form of the panel is natural due to four factors, which have consequences for the number of observations for each player and team:

1. Teams are relegated from the top league and promoted from lower leagues. The impact is that teams and the players will be omitted from the respective seasons until promoted.
2. Players may be transferred to top leagues from lower leagues and vice versa, omitting them from some seasons.
3. Emerging from youth squads and retiring from old age will result in missing values.
4. Data is gathered on a game-by-game basis. Players who do not participate in a game for any given reasons will not receive credit for that game.

Data in football is collected by data providers such as OPTA and Prozone. However, public providers, such as Whoscored, are OPTA-powered websites where data is possible to extract for the public.

We extracted data from Whoscored. Unfortunately, the data does not capture the most essential aspects of dynamic sports. To obtain large quantities of data requires substantial payments, which was not a viable option. Manual data collection was time consuming, so we turned to programming. To obtain the data we developed parsing scripts in R-studio and ran test files (PHP scripts) with Selenium.

5.1 Descriptives

Table 1 depicts summary statistics on the variables applied in our work. It shows basic statistics from 2012/2013-2016/2017 for all matches played in the top five leagues. All variables, except home-/away goal difference have a value greater than, or equal to zero. This is intuitive because a team cannot have negative number of contributions. The variables HomeGoals, AwayGoals, AccurateThroughBallHome and AccurateThroughballAway, have significantly smaller means than other variables. This tells us that these events are the least frequent in each match, which is accurate in terms of how football matches are played.

HomeGoaldiff is positive while AwayGoaldiff is negative signaling that teams playing at home have higher chance of scoring goals at home than away.

Table 1					
Variable	N	Mean	SD	Min	Max
HomeGoals	9128	1.554448	1.316428	0	10
AwayGoals	9128	1.173203	1.157441	0	9
Homegoaldiff	9128	.3812445	1.809222	9	8
Awaygoaldiff	9128	-.3812445	1.809222	8	9
shotsTotalHome	9128	14.04119	5.257566	0	43
shotOnTargetHome	9128	4.871275	2.590569	0	18
keyPassTotalHome	9128	10.37259	4.355867	0	36
dribbleTotalHome	9128	9.473598	4.668684	0	35
foulGivenHome	9128	13.13475	4.277186	0	32
offsideGivenHome	9128	2.438979	1.866358	0	14
dispossessedHome	9128	11.42068	4.598061	0	39
turnoverHome	9128	11.74387	4.185089	0	33
tackleWonTotalHome	9128	19.40809	5.763603	0	48
interceptionAllHome	9128	16.2913	6.062911	0	45
clearanceTotalHome	9128	23.40469	10.54026	0	75
shotBlockedHome	9128	2.59794	1.944458	0	13
foulCommittedHome	9128	13.41937	4.32893	0	33
totalPassesHome	9128	441.0225	115.9043	0	972
TotalCrossHome	9128	22.18788	9.175252	0	81
CrossAccurateHome	9128	5.17167	2.932534	0	21
LongPassesHome	9128	63.28473	13.70601	0	119
AccurateLongPassesTotalHome	9128	33.05719	10.6045	0	91
ThroughBallTotalHome	9128	1.496056	1.99451	0	18
AccurateThroughballHome	9128	.6689308	1.03354	0	11
shotsTotalAway	9128	11.46801	4.70649	0	37
shotOnTargetAway	9128	3.968668	2.313849	0	15
keyPassTotalAway	9128	8.518405	3.883689	0	31
dribbleTotalAway	9128	8.854952	4.574163	0	44
foulGivenAway	9128	12.8142	4.248742	0	32
offsideGivenAway	9128	2.204207	1.773135	0	12
dispossessedAway	9128	11.53298	4.5951	0	36
turnoverAway	9128	11.66827	4.155406	0	34
tackleWonTotalAway	9128	19.6984	5.83124	0	45
interceptionAllAway	9128	16.50646	6.154236	0	47
clearanceTotalAway	9128	27.36624	11.64935	0	95
shotBlockedAway	9128	3.197415	2.182991	0	19
foulCommittedAway	9128	13.81792	4.349065	0	32
totalPassesAway	9128	421.2628	111.6903	0	1081
CrossTotalAway	9128	17.5986	7.637217	0	55
AccurateCrossAway	9128	3.930543	2.451412	0	18
LongPassesAway	9128	63.6755	13.46744	0	114
AccurateLongPassesTotalAway	9128	31.0837	9.847974	0	89
ThroughBallTotalAway	9128	1.33096	1.790959	0	17
AccurateThroughballAway	9128	.5884093	.93263	0	8

Table 1: Variable descriptions of player contributions in Appendix 1

The transfer fees in table 2 are positively skewed and leptokurtic (Appendix 2). Transfer fees lie around 10-20 million euro, with outliers such as 105 million. Estimation of ln_price price is done to remove heteroscedasticity.

Table 2						
Variable	Obs	Mean	Std. Dev.	Min	Max	N
PRICEEuro	1,185	9650949	11800000	150000	105000000	-
Age	1,185	24.7038	3.403723	18	36	-
Africa	1,185	.0919831	.2891241	0	1	109
SouthAmerica	1,185	.1831224	.3869301	0	1	217
Australia	1,185	.0033755	.0580256	0	1	4
EU	1,185	.5721519	.4949756	0	1	678
RestofEurope	1,185	.0464135	.2104677	0	1	55
UK	1,185	.0852321	.2793446	0	1	101
NorthAmerica	1,185	.0059072	.0766631	0	1	7
Asia	1,185	.0118143	.1080955	0	1	14
PremierLeague	1,185	.3510549	.4775016	0	1	416
Laliga	1,185	.164557	.3709368	0	1	195
Serie_A	1,185	.2227848	.4162907	0	1	264
Bundesliga	1,185	.1620253	.3686296	0	1	192
Ligue_1	1,185	.0995781	.299563	0	1	118
Age2	1,185	621.8532	171.692	324	1296	-
Ln_price	1,185	15.52686	1.087006	11.91839	18.46947	-
S_2016	1,185	.2202532	.4145923	0	1	261
W_2016	1,185	.049789	.2176006	0	1	59
S_2015	1,185	.178903	.3834329	0	1	212
W_2015	1,185	.0320675	.1762538	0	1	38
S_2014	1,185	.1611814	.3678535	0	1	191
W_2014	1,185	.028692	.1670099	0	1	34
S_2013	1,185	.1611814	.3678535	0	1	191
W_2013	1,185	.043038	.2030283	0	1	51
S_2012	1,185	.1248945	.3307388	0	1	148
Goalkeeper	1,185	.0413502	.199183	0	1	49
Defender	1,185	.2649789	.4415083	0	1	314
Midfielder	1,185	.4075949	.4915946	0	1	483
Attacker	1,185	.2860759	.4521161	0	1	339
Winter_Transfers	1,185	.1535865	.3607041	0	1	182
Summer_Transfers	1,185	.8464135	.3607041	0	1	1003

Table 2: Summary descriptive of transfer fee, player- and market characteristics

Observations are related to transfers within the same period as player contributions. Nationality, league, transfer window and position is calculated based on dummy variables, hence the number of observations is not very intuitively represented. Nationalities are split into Africa (N=109), North America (7), South America (N=217), Australia (N=4), EU (N=678), rest of Europe (N=55), UK (N=101), Asia (N=14). EU include all countries whom are EU-members, including Schengen, while the rest is self-explanatory.

The descriptive statistics are in line with our prior knowledge of the transfer market: There is a significant difference between transfers completed during the winter and summer windows, 182 and 1003 respectively. Few transfers are done during the winter window due to low supply of high quality players, but the few transfers completed are often due to lack of squad depth and necessity. A note to make is that the English Premier League has the highest number of incoming transfers during our sample period.

6 Methodology

The methodology in this thesis is inspired by the works of (Sæbø, 2016), (McHale et al., 2012) and (S. Dobson & Gerrard, 1999). Initially we test the relationship between transfer fees and player characteristics. Our work will then proceed to create regression based player ratings capturing on-field attacking- and defensive contributions. The final model will be extended to investigate the relationship between real transfer fees, on-field contributions and player characteristics. In formulating this model, we will hopefully be able to determine if the transfer market is efficient.

6.1 Model selection

Expert advice was sought from Thomas Berntsen, sports director in Sarpsborg 08 (a top tier Norwegian football club), for guidance and validation of variables. From the depth interview, conducted in Sarpsborg 08´s offices on 07.08.2017, valuable lessons were learned. All variables obtained are relevant in formulating a rating system. However, he mentioned that the four pillars he applies to identify quality players for Sarpsborg 08 are harder to quantify with our data:

1. Ability and willingness to train (mentality and fitness).
2. Team spirit on- and off pitch (attitude and culture).
3. Physical ability (Stamina, speed, strength).
4. Ability to understand the game (Intelligence).

These "pillars" are specific to Sarpsborg 08, meaning that clubs are heterogeneous in valuating football players. Implying that the value a club will attach to a given player will differ across bidders. This means more often than not, transfers in the market will be a non-zero sum game as explained by Von Neumann (2004). In the interview, Berntsen emphasized that defenders are the hardest to assess, as many positive contributions are not quantifiable (reading the game, man-marking, positioning).

Our datasets are unbalanced panel data and model specification leads us towards the choice of pooled OLS, random effects (RE) or fixed effect (FE) model (Brooks, 2014). The Fixed effects model implies that there are heterogeneous, time invariant effects μ_i that are specific to each player, and correlated with the independent variables. Applying a random effects model without controlling for fixed effects through the Hausman test, would lead to serial correlation in the error term and omitted variable bias. The effect of this would mean that the error terms are not i.i.d because the fixed effects would be captured in ϵ_i.

Random effects:

$$y_{it} = \alpha + \beta x_{it} + \omega_{it}, \qquad \omega_{it} = \epsilon_i + v_{it}$$

Fixed Effects:

$$y_{it} = \alpha + \beta x_{it} + u_{it}, \qquad u_{it} = \mu_i + v_{it}$$

The Hausman test shows that we keep the null (Appendix 3): Fixed effect method is appropriate. For additional testing, we ran the Breusch pagan test for pooled OLS vs Random effects, and the Fixed effect method was again confirmed as appropriate. Our dataset therefore has unobservable effects, μ_i, that are correlated with explanatory variables. The general fixed effects model:

$$Y_{it} = \alpha_1 + \beta_2 X_{it} + \cdots + \beta_T X_{iT} + \mu_i + v_{iT}$$

$$\overline{Y_i} = \alpha_1 + \beta_2 \bar{X}_{2i} + \cdots + \beta_T \bar{X}_{Ti} + \mu_i + \bar{v}_i$$

$$Y_{it} - \overline{Y_i} = \beta_2 (X_{it} - \bar{X}_{2i}) + \cdots + \beta_T (X_{iT} - \bar{X}_{Ti}) + v_{iT} - \bar{v}_i$$

FE controls for a lot of potential omitted variable bias, and the demeaning has the positive implication of eliminating individual specific effects. For our research, individual specific effects will refer to which team each player plays for. Eliminating team specific effects, enable us to create an objective rating by comparing strong teams and weak teams on the same grounds. This will hopefully give a neutral ground for comparison of player performance. A drawback is that FE limits what we can estimate, as estimation of time invariant effects and out of sample inference are not possible.

According to Pollard (1985), goals are a random result of shots. Implicitly stating that match outcomes are explained by randomness, which makes it a parsimonious predictor for rating football players. We believe that contributions leading to shot opportunities, explained in figure 3, says more about the quality of a player than random goals and assists.

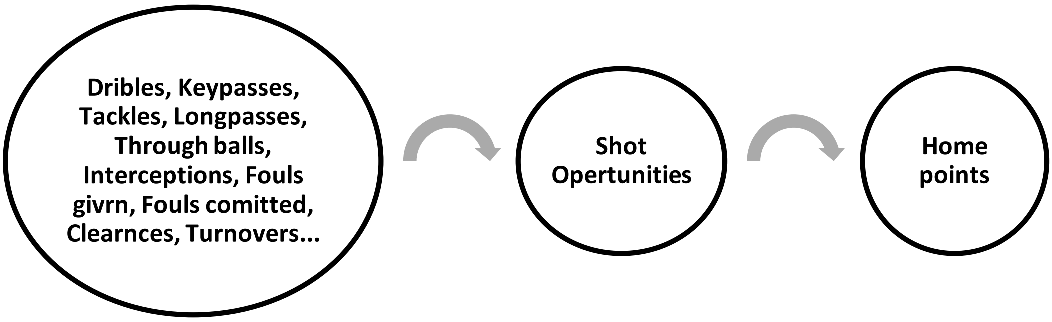


Figure 4: Skill modelling: adopted from (McHale et al., 2012)

Most studies within the field has shown that OLS is appropriate for estimation of transfer fees and impact of characteristics (Frick, 2007). We are highly interested in time invariant factors such as, nationality/League/position, and unobservable individual specific effects (innate ability). When conducting fixed effects method these variables would be omitted, hence the appropriate and chosen method for model two and three is OLS.

6.2 Econometric method

Model 1 – The Rating model (FE):

The total contribution of a team is the sum of individual contributions. Based on this we may obtain our beta coefficients by regressing on total team contribution. We are interested in these coefficients as they can be interpreted as a players' marginal contribution in creating shot opportunities for his team. Positive contributions will hence increase team expected shot, and vice versa for negative contributions. Estimation in this manner allows us to carry estimated coefficients from team over to players, which give the following regression:

$$E[Shot\ opportunity_H] = \alpha_0 + \beta_1 X1_{it} + \beta_2 X2_{it} + \cdots + \beta_N XN_{it}$$

$$E\left[S_{H_i}\right] = F(H_{x1} + \cdots + H_{xn} + A_{x1} + \cdots + A_{xn})$$

The expected shot opportunities created by a given team is a function of home and away contributions in each game. As we model from the perspective of the home team, negative away contributions such as Ax, will capture the defensive capabilities home team players. Negative away contributions are how we model the defensive rating $E[Def_i]$. The model could be switched around to estimate expected shot for away.

$$E[reduction\ of\ away\ team\ shot\ opportunity] = E[Deff_i]$$

$$I_{Rating_i} = E\left[S_{H_i}\right] + E[Def_i]$$

See appendix 1 for detailed descriptions of notation

The final rating will be the sum of offensive and defensive contributions with respect to the increase of expected shot for home team. The contributions will be normalized to give intuitive results.

<u>Model 2 – The transfer fee model (OLS):</u>

Log of transfer fees will be regressed against player specific characteristics to see which are significant. Age is included because with experience the transfer value is expected to increase. Age squared is included to correct for the fact that at some point in time, football players will start to become less valuable as they become. Nationality and position are player specific dummies, while league dummies refer to the buying clubs league.

Transferyear dummies are present to capture time effects in different in transfer windows. The model to be tested is the following:

$$logPRICEEuro = age + age^2 + LeagueDummies + NationalityDummies + PositionDummies + TransferyearDummies$$

Model 3: The Augmented Transfer fee model (OLS):

Model three will be an augmentation of model two, which includes the objective player rating estimated in model one. This will hopefully increase the explanatory power of the transfer fee model. If this is true, then our rating has an ability to identify quality players, and that investors are somewhat rational in the market. Diagnostics test will be conducted, and results will show in appendix 6.

$$logTransferFee = age + age^2 + LeagueDummies + NationalityDummies + PositionDummies + TransferyearDummies + I_{Rating_i}$$

6.3 Model expectations

Our expectations for the model is as following, and any major deviation will suggest inefficiency in the market:

- Transfer fees are highly correlated with age^2. and age. Age^2 should have a negative coefficient, because when a players age increases, transfer fee will start to drop at some point. Older players are in general more expensive than younger players with less experience, leading us to expect positive and significant age coefficient.
- Positive player contributions on the pitch should be reflected positively in the price. Higher rating should increase transfer fee, meaning our model identifies quality players and investors are somewhat rational.
- We expect that nationality should not have a significant effect, i.e. there should be no difference in the price of two similarly rated players from different parts of the world.
- Midfielders will be overestimated because our rating is based on both offensive and defensive contributions.
- We expect that investors pay a premium for attackers and midfielders, due to their popularity and goal scoring ability.

7 Analysis

7.1 The player rating model

For the rating model, we have cut insignificant explanatory factors for practical purposes; the full list may be found in appendix 4. Variables in table four represent significant individual contributions that increase or decrease the expected shot for home team. The rating model will represent player i´s increase in contributions in respect to affecting his teams expected shot opportunity. For instance, if a given player has one key pass he will increase expected shot opportunities for his team by 1.038094, and the opposite intuition holds for variables with negative coefficients. We include variables with degree of significance of ten, five and one percent based on observations for home and away team. The variable coefficients are obtained from the all team-specific data, and are applied for rating each individual player.

R^2: $Within = 0.8265;\ Between = 0.9467; Overall = 0.8487$

The Rating Model-Significant Factors				
shotsTotalHome	Coef.	Std.Err.	t	P>t
keyPassTotalHome	1.038094	.0061062	170.01	0.000***
dribbleTotalHome	.0257492	.0052436	4.91	0.000***
foulGivenHome	.0523754	.0055724	9.40	0.000***
offsideGivenHome	-.0383464	.0116011	-3.31	0.001***
turnoverHome	-.0122321	.0053896	-2.27	0.023**
clearanceTotalHome	-.0049375	.0022888	-2.16	0.031**
shotBlockedHome	-.0260977	.0134291	-1.94	0.052*
foulCommittedHome	-.0179812	.0055905	-3.22	0.001***
passCrossTotalHome	.0470369	.0029324	16.04	0.000***
passLongBallTotalHome	-.0117116	.002526	-4.64	0.000***
passLongBallAccurateHome	.0070936	.0031725	2.24	0.025**
shotOnTargetAway	.0322043	.0123454	2.61	0.009***
keyPassTotalAway	-.0344465	.0084817	-4.06	0.000***
tackleWonTotalAway	-.0073012	.0039332	-1.86	0.063*
interceptionAllAway	-.0092711	.0036607	-2.53	0.011**
passLongBallAccurateAway	-.0091802	.0023668	-3.88	0.000***
_cons	3.214274	.2434752	13.20	0.000

Table 3: Results of player rating model

The model is a good fit ($R^2 = 0.8487$) for modelling how shot opportunities (shotsTotalHome) are created. Explanatory power of the model is naturally extremely high as these variables explain important events in the game. Key passes are defined as passes directly leading to a shot opportunity, and is the most significant variable with a t-value of 170.01.

Coefficients signs are intuitive understanding of what increases and decreases opportunities on goal. For instance, if the home team have many fouls committed and turnovers, they will lose possession and hindering the home team in creating opportunities. Clearances and shots blocked reduces shot opportunities because the away team is applying pressure. Total long balls from home team have negative coefficients, but accurate long balls are positive. Away team coefficients signs explain the opposite of home team coefficients signs. Key pass, shots on target by away and accurate long balls away will reduce the shot opportunities for the home team.

We model offensive contributions based on the home team, and defensive contributions based on away team. This is because both have the same objective; to reduce the opponents shot opportunities and increasing their own. As table 4 depicts, home team tackles and interceptions are insignificant when modelling home team shot opportunities. Home team tackles and interceptions are variables related to their defensive objective. Defensive contributions for a given player is captured by the negative away team coefficients.

The ranking of players in table 4 is done through the normalization defensive and offensive contribution. We are very pleased with our results, as it captures top players in Europe, by portraying highly regarded players. During the timespan of our dataset, several of our top thirty players have been contenders and winners, of Ballon d´or - the most prestigious individual award in European football.

Table 4				
Rank	PlayerID	Rating	No. Matches	Position
1	Toni Kroos	1	79	Midfielder
2	Lionel Messi	.9997755	83	Attacker
3	Roberto Trashorras	.9993654	71	Midfielder
4	Marek Hamsik	.996997	89	Midfielder
5	Cristiano Ronaldo	.9968117	85	Attacker
6	Johannes Geis	.9939967	63	Midfielder
7	Cesc Fàbregas	.9936093	80	Midfielder
8	Kevin De Bruyne	.9929601	76	Midfielder
9	Daniel Parejo	.9907831	78	Midfielder
10	Radja Nainggolan	.9900004	87	Midfielder
11	Ivan Rakitic	.9896111	85	Midfielder
12	Zlatan Ibrahimovic	.9893494	79	Attacker
13	Paul Pogba	.9892403	83	Midfielder
14	Luis Suárez	.9888618	80	Attacker
15	Arturo Vidal	.9886413	73	Midfielder
16	Mark Noble	.9881441	82	Midfielder
17	Miralem Pjanic	.9881024	79	Midfielder
18	Gabi	.9880582	90	Midfielder
19	Luca Cigarini	.9873828	63	Midfielder
20	Dimitri Payet	.987341	88	Midfielder
21	Daniel Baier	.9859157	78	Midfielder
22	Antonio Candreva	.985635	90	Midfielder
23	Beñat Etxebarria	.9850889	77	Midfielder
24	Steven Gerrard	.984543	50	Midfielder
25	Ryad Boudebouz	.9845006	86	Attacker
26	Borja Valero	.9844887	83	Midfielder
27	Max Kruse	.9833549	77	Attacker
28	Eden Hazard	.9826443	88	Attacker
29	Yaya Touré	.9815438	78	Midfielder
30	Wayne Rooney	.980732	71	Attacker

Table 4: Player ranking

Our top thirty ranking include twenty-one midfielders and eight attackers with no goalkeepers and defenders. This confirms our suspicion that the rating would overestimate midfielders due to their defensive and offensive contributions. The maximum number of possible games played are ninety in this dataset, apart from players in Bundesliga (eighty-five games). This is because we model only based on home games.

Further on, we see that most players included in the table are consistent starting players, as they are close to ninety played games. The surprising result of Steven

Gerrard (fifty games), Luca Cigarini (sixty-three games) and Johannes Geis (sixty-three games) are worth mentioning. These players have probably performed significantly better than similar players during their games, suggestion that they are more effective.

Most highly rated players are between twenty-five-thirty years old, which is considered a footballer's career peak. There are some exceptions, with Johannes Geis in sixth- and Paul Pogba in thirteenth place. These players have not yet reached their prime years in football. At the start of our dataset these players were nineteen years old. During the last five years, they have become consistent starters for their teams and are found amongst the top thirty in our model. They have experience from a young age, good performance and high future potential. These types of players outperform their peers and experienced (older) players, which might indirectly suggest a premium price for such players.

For the next model, we will include the player rating to see if we can explain more of the variation in transfer fees.

7.2 The transfer fee model

The explanatory power of the model is somewhat lower than expected, but we still have interesting results. The premier league is the only league which have a premium for inbound transfers, with a significant positive effect. This can be explained by the new television rights of 5 billion; teams in the premier league are, on average, richer than other clubs, who require a premium when selling to the Premier League.

$R^2 = 0.1909$

Table 5				
Ln_price	Coef.	Std. Err.	t	P>t
Age	.5714324	.1031452	5.54	0.000***
Africa	.5639787	.3894411	1.45	0.148
SouthAmerica	.9727457	.3850358	2.53	0.012**
Australia	-.5766509	.6243518	-0.92	0.356
EU	.6300481	.3806473	1.66	0.098*
RestofEurope	.413239	.4023778	1.03	0.305
UK	.3744055	.3885281	0.96	0.335
NorthAmerica	0	(omitted)		
Asia	.2569924	.461121	0.56	0.577
PremierLeauge	.7597053	.0919839	8.26	0.000***
Laliga	0	(omitted)		
Serie_A	-.0037815	.0942604	-0.04	0.968
Bundesliga	.1156831	.1025138	1.13	0.259
Ligue_1	.0047963	.1175046	0.04	0.967
Age2	-.0117251	.0020473	-5.73	0.000***
S_2012	.0584063	.1806268	0.32	0.746
W_2013	-.4613593	.2126944	-2.17	0.030**
S_2013	.059903	.1767695	0.34	0.735
W_2014	-.3527107	.2343985	-1.50	0.133
S_2014	.1737246	.1771056	0.98	0.327
W_2015	0	(omitted)		
S_2015	.3239412	.1756002	1.84	0.065*
W_2016	-.4440999	.2063588	-2.15	0.032**
S_2016	.2353896	.1736364	1.36	0.175
Attacker	.3378078	.1550615	2.18	0.030**
Midfielder	.268702	.1531755	1.75	0.080*
Defender	-.0244303	.1554466	-0.16	0.875
Goalkeeper	0	(omitted)		
_cons	7.471518	1.338718	5.58	0.000

Table 5: The transfer fee model

Age is significantly positive as expected; as players' experience increase the price of players increase. Similarly, Age2 is negatively significant on the one percent level, controlling for diminishing skills of older players. Nationalities coefficients are insignificant for all nationalities on the one and five percent level, except for South American players, which have a large positive premium. This may refer to biases within the transfer market, as being South American automatically increases transfer fee. The racial bias has been detected studies as well (Frick, 2007, p. 15).

Seasonality has significantly negative effects during the winter windows of 2013 and 2016. There are two possible explanations that may explain these results. Transfer activity is low during the winter window. Secondly, players are reluctant to transfer before large championships, as the European Championship was played during the proceeding summer of 2016. Transferring to a new team might lead to less playing time, and players are in general reluctant in risking call-up to their respective national teams.

Attackers are the only significant positions variable. This may not come as a shock, as goals scored is the most valuable aspect of any game. Attackers are in general more popular, and often require premiums (Neymar sold for 220 million), and are the superstars of the game.

The augmented model is identical to the one described in table five, but includes the player specific ratings formulated in 6.1. In table six we have cut out insignificant variables for practical purposes. Table including insignificant variables may be found in appendix 4.

$R^2 = 0.2969$

Table 6

Ln_price	Coef.	Std.Err	t	P>t
Age	.2486808	.0967784	2.57	0.010***
SouthAmerica	.4477974	.0702415	6.38	0.000***
PremierLeauge	.6583511	.056829	11.58	0.000***
Age2	-.0060401	.0019101	-3.16	0.002***
W_2013	-.4048322	.1325128	-3.06	0.002***
S_2016	.3084086	.0659736	4.67	0.000***
Midfielder	-.215459	.0699747	-3.08	0.002***
Defender	-.2493918	.0643863	-3.87	0.000***
I_Rating1	1.8154	.1153884	15.73	0.000***
_cons	11.81129	1.193513	9.90	0.000***

Table 6: The augmented transfer fee model

The new model has higher explanatory power than the previous. The inclusion of player rating shows that the model captures some of the variation in transfer fees, suggesting that our model can identify quality players.

Variables age, age^2, South America, Premier League and W_2013 are still significant with same signs. Midfielder has a significant negative coefficient. This can be interpreted as a correction as our rating model overstates the impact of midfielders. The variable Defenders is now significant with a negative coefficient. The reasoning may be that Defenders are harder to valuate, because lack of available variables capturing defensive skill as mentioned in the interview with Thomas Berntsen. W_2013 is negative on the one percent level. Players transferred during this period had a price reduction, but explanation for this cannot be established. S_2016 is positively significant on the one percent level. The European Championship in France was played during the summer of 2016, which can inflate the price of players who did well.

By our model, transfer fee of any given player in the market is highly defined by the following:

1. Rating (t-value=15.73): Highly rated players are positively linked to higher transfer fees.
2. English Premier League (t-value=11.58): If any given player is transferred to the Premier league, he will have a price which is 68 percent higher than if he went to any other leagues. This follows our expectation that leagues with more capital tend to pay premium for players.
3. South American players (t-value=6.38): Our findings show that a premium exists on South American players in the transfer market. South American players are automatically 44.8 percent more expensive than others. From this we can partly state that there exists some degree of inefficiency in the transfer market. In theory, teams could sell their South American players and acquire players of similar rating but difference nationality and make a profit.

8 Conclusion

Our thesis seeks to test whether the transfer market is efficient or not. Model one has high explanatory power, but lacks variables explaining team dynamics, quality of chances created. Model two and three lack important off-pitch variables such as wages, injury proneness, contract length and marketability. However, most of our findings were still consistent with previous studies within the field. Both age variables are highly significant in the determining prices in both models, as well to nationality. One of the more interesting findings were that the player rating model was most significance in explaining transfer fees.

We believe that the limitations are the basis for the low explanatory power of the transfer fee models. The main finding answering the question is the premium of South American players, which coincides with earlier research. Since nationality is a strong determinant of price, means that there is racial biasness in the market. This is the only basis for deeming market inefficiency. Nationality should not be a measure of quality in any market, leading us to conclude that the transfer market is **weakly inefficient.**

Studying the transfer market and football players as assets has been an intriguing project. We hope that our thesis will encourage peers to move out of their comfort zone by analyzing untraditional areas. Further studies related to ours should model the rating after expected goals and assists. These metrics consider how and where these events were created. In our best effort, we attempted to perform this analysis, but came short in the data collection process. In regards to the transfer fee models, the abovementioned off-pitch elements should be included, especially marketability.

9 Bibliography

Big 5 - Transfer Window Analysis Summer 2016. (2016). Retrieved from https://www.fifatms.com/data-reports/reports/

Brooks, C. (2014). *Introductory econometrics for finance* (3rd ed. ed.). Cambridge: Cambridge University Press.

Carmichael, F., Forrest, D., & Simmons, R. (1999). The Labour Market in Association Football: Who Gets Transferred and for How Much? *Bulletin of Economic Research, 51*(2), 125-150. doi:10.1111/1467-8586.00075

Carmichael, F., & Thomas, D. (1993). Bargaining in the transfer market: theory and evidence. *Applied Economics, 25*(12), 1467-1476. doi:10.1080/00036849300000150

Dawson, P., Dobson, S., & Gerrard, B. (2000). Estimating Coaching Efficiency in Professional Team Sports: Evidence from English Association Football. *Scottish Journal of Political Economy, 47*(4), 399-421. doi:10.1111/1467-9485.00170

Dobson, S., & Gerrard, B. (1999). The determination of player transfer fees in English professional soccer. *Journal of Sport Management, 13*(4), 259-279.

Dobson, S., & Goddard, J. (2001). *The Economics of Football*. Cambridge, UNKNOWN: Cambridge University Press.

Fama, E. F. (1970). Efficient Capital Markets - Review of Theory and Empirical Work. *Journal* *of Finance, 25*(2), 383-423. doi:Doi 10.2307/2325486

Frick, B. (2007). The football players' labor market: Empirical evidence from the major European leagues. *Scottish Journal of Political Economy, 54*(3), 422-446. doi:DOI 10.1111/j.1467-9485.2007.00423.x

Gerrard, B. (2007). Is the Moneyball approach transferable to complex invasion team sports? *International Journal of Sport Finance, 2*(4), 214-230.

Leach, S., & Szymanski, S. (2015). Making Money Out of Football. *Scottish Journal of Political Economy, 62*(1), 25-50. doi:10.1111/sjpe.12065

McHale, I. G., Scarf, P. A., & Folker, D. E. (2012). On the Development of a Soccer Player Performance Rating System for the English Premier League. *Interfaces, 42*(4), 339-351. doi:10.1287/inte.1110.0589

Nash, J. F. (1950). The Bargaining Problem. *Econometrica, 18*(2), 155-162. doi:10.2307/1907266

Peeters, T., & Szymanski, S. (2014). Financial fair play in European football. *Economic Policy, 29*(78), 343-390. doi:10.1111/1468-0327.12031

Pollard, R. (1985). 69.9 Goal-Scoring and the Negative Binomial Distribution. *The Mathematical Gazette, 69*(447), 45-47. doi:10.2307/3616453

Sloane, P. (1969). THE LABOUR MARKET IN PROFESSIONAL FOOTBALL*. *British Journal of Industrial Relations, 7*(2), 181-199. doi:10.1111/j.1467-8543.1969.tb00560.x

Sloane, P. J. (1971). Economics of Professional Football - Football Club as a Utility Maximiser. *Scottish Journal of Political Economy, 18*(2), 121-146. doi:DOI 10.1111/j.1467-9485.1971.tb00979.x

Sloane, P. J. (2015). The Economics of Professional Football Revisited. *Scottish Journal of Political Economy, 62*(1), 1-7. doi:10.1111/sjpe.12063

Sæbø, O. D. (2016). *Evaluating the efficiency of the association football transfer market using regression based player ratings* (Vol. 2015:79). Molde: Molde University College, Specialized University in Logistics.

Tunaru, R., Clark, E., & Viney, H. (2005). An option pricing framework for valuation of football players. *Review of Financial Economics, 14*(3), 281-295. doi:http://dx.doi.org/10.1016/j.rfe.2004.11.002

Von Neumann, J. (2004). *Theory of games and economic behavior* (60th anniversary ed. ed.). Princeton, N.J: Princeton University Press.

10 Appendix

Appendix 1: Variables description

keyPassTotalHome: Passes leading shots at goal, for home team.

dribbleTotalHome: Total succesfull dribbles, for home team

foulGivenHome: Foul given to home team

offsideGivenHome: Home Team is couthg in offside

dispossessedHome: Home team is disposessd of the balll, without attemting to dribble.

turnoverHome: Loss of possesien due to a mistake.

tackleWonTotalHome: Total tackles won, for home team.

interceptionAllHome: Preventing opponents pass from reaching their intended target.

clearanceTotalHome: An action by a defender that temporarily removes attacking threat, for home team.

shotBlockedHome: Prevention of by an outfiels player of an opponents shot reaching the goal

foulCommittedHome: An illeagal manoeuvre by a player that results in a free kick for opposing team

totalPassesHome: All passes. (Failed & Succesful)

passCrossTotalHome: All Crosses. (Failed & Succesful)

passCrossAccurateHome: Acurate crosses. (Succesful)

passLongBallTotalHome: All long passes. (Failed & Succesful)

passLongBallAccurateHome: Acurate long passes. (Succesful)

passThroughBallTotalHome: All passes between opposition players in their defensive line. (Failed & Succesful)

passThroughBallAccurateHome: passes between opposition players in their defensive lin (Failed & Succesful)

All definitions are the same the away team, with notation "away".

Appendix 2: Test for skewness and kurtosis.

```
sktest PRICEEuro
```

Skewness/Kurtosis tests for Normality

Variable	Obs	Pr(Skewness)	Pr(Kurtosis)	adj chi2(2) (joint)	Prob>chi2 (joint)
PRICEEuro	1,185	0.0000	0.0000	.	0.0000

Appendix 3: Hausman test

Hausman test Fixed effect v Random effects

	Coefficients			
	(b) fixed	(B) random	(b-B) Difference	sqrt(diag(V_b-V_B)) S.E.
keyPassTotalHome	.9478065	.9534568	-.0056503	.0011464
dribbleTotalHome	.0254718	.0260454	-.0005736	.0010691
foulGivenHome	.0335727	.0338677	-.000295	.0020123
offsideGivenHome	-.0309288	-.027972	-.0029568	.0014332
dispossessedHome	.0019085	-.0012428	.0031513	.0013016
turnoverHome	-.0217764	-.0218001	.0000237	.0008715
tackleWonTotalHome	-.0027095	-.0019538	-.0007557	.0008283
interceptionAllHome	-.0018448	-.0004035	-.0014413	.0006529
clearanceTotalHome	-.0101019	-.0115075	.0014056	.0005408
shotBlockedHome	-.1394444	-.1437301	.0042857	.0018547
foulCommittedHome	-.0569451	-.0544159	-.0025292	.0020326
totalPassesHome	-.000324	-.0003816	.0000576	.0001078
TotalCrossHome	.0197285	.0170231	.0027054	.0009225
CrossAccurateHome	.0015651	.0001343	.0014309	.0013199
LongPassesHome	-.0218357	-.0205422	-.0012935	.0006671
AccurateLongPassesTotalHome	.018469	.0194652	-.0009962	.0006354
ThroughBallTotalHome	.000315	-.0023647	.0026797	.0036033
AccurateThroughballHome	.0435831	.0468548	-.0032718	.0036644
shotsTotalAway	.1570449	.1602146	-.0031697	.0013125
shotOnTargetAway	-.0442605	-.0442473	-.0000132	.001208
keyPassTotalAway	-.1093256	-.1111669	.0018413	.0011798
dribbleTotalAway	-.0055345	-.0048289	-.0007057	.0009203
foulGivenAway	.0341392	.0362808	-.0021416	.0020499
offsideGivenAway	.0276766	.0257528	.0019237	.0023126
dispossessedAway	.0109208	.010613	.0003078	.0009654
turnoverAway	.0115056	.0106217	.000884	.0008658
tackleWonTotalAway	-.014974	-.0126108	-.0023633	.0008618
interceptionAllAway	-.0127314	-.0118557	-.0008757	.0005296
clearanceTotalAway	.0126487	.0127078	-.000059	.0005165
shotBlockedAway	.3766306	.3764173	.0002133	.0014294
foulCommittedAway	.0161833	.0162087	-.0000255	.0018336
totalPassesAway	-.0003119	-.00033	.0000181	.0000557
CrossTotalAway	-.0021627	-.0032255	.0010628	.0008279
AccurateCrossAway	-.0141288	-.0137777	-.0003511	.0011834
LongPassesAway	.0134307	.0135631	-.0001325	.0005569
AccurateLongPassesTotalAway	-.0176492	-.0171604	-.0004887	.0006429
ThroughBallTotalAway	.0057496	.0030457	.002704	.0032106
AccurateThroughballAway	-.0431216	-.0453385	.0022169	.0035613

b = consistent under Ho and Ha; obtained from xtreg
B = inconsistent under Ha, efficient under Ho; obtained from xtreg

Test: Ho: difference in coefficients not systematic

chi2(38) = (b-B)'[(V_b-V_B)^(-1)](b-B)
= 79.57
Prob>chi2 = 0.0001

Appendix 4: Rating Model including insignificant variables.

The Rating Model-Complete Table				
shotsTotalHome	Coef.	Std.Err.	t	P>t
keyPassTotalHome	.9478065	.006975	135.89	0.000***
dribbleTotalHome	.0254718	.0051189	4.98	0.000***
foulGivenHome	.0335727	.018655	1.80	0.072*
offsideGivenHome	-.0309288	.0109962	-2.81	0.005***
dispossessedHome	.0019085	.0063815	0.30	0.765
turnoverHome	-.0217764	.0053779	-4.05	0.000***
tackleWonTotalHome	-.0027095	.005415	-0.50	0.617
interceptionAllHome	-.0018448	.0037118	-0.50	0.619
clearanceTotalHome	-.0101019	.0026065	-3.88	0.000***
shotBlockedHome	-.1394444	.013925	-10.01	0.000***
foulCommittedHome	-.0569451	.0186247	-3.06	0.002***
totalPassesHome	-.000324	.0002977	-1.09	0.276
passCrossTotalHome	.0197285	.0040174	4.91	0.000***
passCrossAccurateHome	.0015651	.0104709	0.15	0.881
passLongBallTotalHome	-.0218357	.0029004	-7.53	0.000***
passLongBallAccurateHome	.018469	.0036551	5.05	0.000***
passThroughBallTotalHome	.000315	.0167806	0.02	0.985
passThroughBallAccurateHome	.0435831	.0302579	1.44	0.150
shotsTotalAway	.1570449	.0128109	12.26	0.000***
shotOnTargetAway	-.0442605	.0129438	-3.42	0.001***
keyPassTotalAway	-.1093256	.013445	-8.13	0.000***
dribbleTotalAway	-.0055345	.005212	-1.06	0.288
foulGivenAway	.0341392	.0189362	1.80	0.071*
offsideGivenAway	.0276766	.0119784	2.31	0.021**
dispossessedAway	.0109208	.006369	1.71	0.086*
turnoverAway	.0115056	.0054108	2.13	0.033**
tackleWonTotalAway	-.014974	.0052714	-2.84	0.005***
interceptionAllAway	-.0127314	.0036331	-3.50	0.000***
clearanceTotalAway	.0126487	.0024497	5.16	0.000***
shotBlockedAway	.3766306	.011404	33.03	0.000***
foulCommittedAway	.0161833	.0183128	0.88	0.377
totalPassesAway	-.0003119	.0002782	-1.12	0.262
passCrossTotalAway	-.0021627	.0044665	-0.48	0.628
passCrossAccurateAway	-.0141288	.0119303	-1.18	0.236
passLongBallTotalAway	.0134307	.0027633	4.86	0.000***
passLongBallAccurateAway	-.0176492	.0036878	-4.79	0.000***
passThroughBallTotalAway	.0057496	.0179977	0.32	0.749
passThroughBallAccurateAway	-.0431216	.0329343	-1.31	0.190
_cons	2.984857	.3184253	9.37	0.000

Appendix 5: Augmented Transfer Fee model including insignificant variables

Table 6 - full list				
Ln_price	Coef.	Std.Err	t	P>t
Age	.2384459	.0974866	2.45	0.015**
SouthAmerica	.4263711	.1096348	3.89	0.000***
Australia	-.3564066	.4688272	-0.76	0.447
EU	-.0157211	.0956249	-0.16	0.869
RestofEurope	-.0846144	.1524595	-0.55	0.579
UK	-.1852114	.1310569	-1.41	0.158
NorthAmerica	-.2114286	.3587271	-0.59	0.556
Asia	-.3973129	.2599636	-1.53	0.127
PremierLeauge	.7191405	.0845837	8.50	0.000***
Laliga	0	(omitted)		
Serie_A	-.0035191	.0866306	-0.04	0.968
Bundesliga	.0982607	.0942234	1.04	0.297
Ligue_1	-.0373272	.1080316	-0.35	0.730
Age2	-.0058656	.0019237	-3.05	0.002***
S_2012	-.0260987	.1661064	-0.16	0.875
W_2013	-.3863721	.1955451	-1.98	0.048**
S_2013	-.107227	.1628616	-0.66	0.510
W_2014	-.2810781	.2154808	-1.30	0.192
S_2014	.0785032	.1628997	0.48	0.630
W_2015	0	(omitted)		
S_2015	.2524438	.1614601	1.56	0.118
W_2016	-.1856253	.1904752	-0.97	0.330
S_2016	.3213787	.1596895	2.01	0.044**
Attacker	-.1574719	.1464706	-1.08	0.283
Midfielder	-.3874288	.1477387	-2.62	0.009***
Defender	-.3550939	.1446384	-2.46	0.014**
Goalkeeper	0	(omitted)		
I_Rating1	1.794143	.1225509	14.64	0.000***
_cons	12.09052	1.219433	9.91	0.000

6. Miscellaneous OLS tests

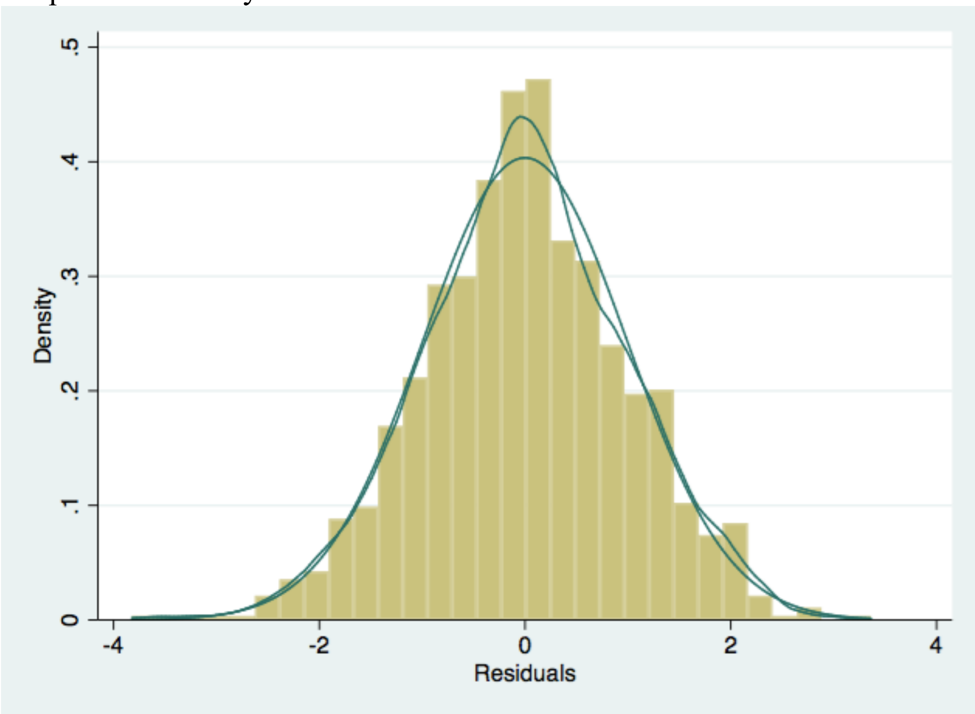
Breusch Pagan test for Heteroscedasticity for OLS.

```
. estat hettest

Breusch-Pagan / Cook-Weisberg test for heteroskedasticity
         Ho: Constant variance
         Variables: fitted values of Ln_price

         chi2(1)      =     1.34
         Prob > chi2  =   0.2466
```

Graph for normality in residuals OLS



JB test for normality (formal)

```
. jb resid
Jarque-Bera normality test:  2.375 Chi(2)  .3049
Jarque-Bera test for Ho: normality:
```

Ramsey test for omitted variable bias

Transfer fee model

```
. ovtest

Ramsey RESET test using powers of the fitted values of Ln_price
       Ho:  model has no omitted variables
                 F(3, 1157) =      0.95
                   Prob > F =      0.4151
```

Augmented transfer fee model.

```
. ovtest

Ramsey RESET test using powers of the fitted values of Ln_price
       Ho:  model has no omitted variables
                F(3, 1156) =      1.65
                  Prob > F =      0.1758
```