Determinants of Soccer Players' Market Values in Major European Leagues*

A Comparative Linear Regression Analysis Across Five Leagues

John Zhang

November 26, 2024

Table of contents

1 Introduction

The global soccer transfer market is a dynamic and high-stakes environment where players' market values are central to team-building strategies, investments, and negotiations. Under-

*Code and data are available at: https://github.com/Clearsky21z/Player_Market_Value_Analysis

standing what drives these values is a critical question for analysts, clubs, and stakeholders. While common factors such as goals scored or club reputation are widely acknowledged to affect valuations, their relative importance likely varies across leagues with differing levels of competition, financial resources, and playing styles.

This paper focuses on five major European soccer leagues—Premier League(England), La Liga (Spain), Serie A (Italy), Bundesliga (Germany), and Ligue 1 (France). These leagues are not only among the most competitive but also vary in terms of audience size, financial backing, and player scouting pipelines, making them ideal for a comparative study of market value determinants.

Using player data from the 2023/24 season, this study develops league-specific linear regression models with market value as the dependent variable. Key predictors include age, goals, assists, club ranking, national team ranking, minutes played, and position. The analysis reveals both shared and league-specific patterns in the factors influencing market value. For instance, goals and assists consistently emerge as significant predictors, but their impact differs by league. Similarly, younger players tend to command higher valuations universally, with age effects most pronounced in Serie A.

These findings contribute to our understanding of how soccer markets function across different leagues and provide actionable insights for clubs and agents. The structure of this paper is as follows: Section 2 describes the dataset and variables, Section 3 outlines the modeling approach, Section 4 presents the results, and Section 5 discusses the implications and limitations of the findings.

2 Data

2.1 Dataset Overview

The data includes information on players from the top five European leagues during the 2023/24 season. Variables of interest are:

- **Market Value**: Player's estimated value in Euros.
- **Age**: Player age at the start of the season.
- **Goals and Assists**: Indicators of offensive contributions.
- **Club Ranking**: Rank of the player's club based on league and international performance.
- **National Team Ranking**: Ranking of the player's national team (1 to 210).
- **Minutes Played**: Total minutes played during the season.
- **Position**: Categorized as Defender (DF), Midfielder (MF), Forward (FW), and hybrids like FWDF.

2.2 Summary Statistics

Table 1 summarizes the descriptive statistics for the Premier League dataset.

Table 1: Summary Statistics for the Premier League Dataset

Variable	value	mean	ranking	team	played	sd	min	max
market	20992333	NA	NA	NA	NA	NA	NA	NA
age	NA	24.79	NA	NA	NA	4.38	16	38
goals	NA	2.24	NA	NA	NA	3.68	0	27
assists	NA	1.64	NA	NA	NA	2.41	0	13
club	NA	NA	130.05	NA	NA	NA	NA	NA
national	NA	NA	NA	17.60	NA	NA	NA	NA
minutes	NA	NA	NA	NA	1310.96	NA	NA	NA
market	23232211	NA	NA	NA	NA	NA	NA	NA
club	NA	NA	193.08	NA	NA	NA	NA	NA
national	NA	NA	NA	22.69	NA	NA	NA	NA
minutes	NA	NA	NA	NA	1001.09	NA	NA	NA
market	100000	NA	NA	NA	NA	NA	NA	NA
club	NA	NA	1.00	NA	NA	NA	NA	NA
national	NA	NA	NA	1.00	NA	NA	NA	NA
minutes	NA	NA	NA	NA	1.00	NA	NA	NA
market	180000000	NA	NA	NA	NA	NA	NA	NA
club	NA	NA	838.00	NA	NA	NA	NA	NA
national	NA	NA	NA	173.00	NA	NA	NA	NA
minutes	NA	NA	NA	NA	3420.00	NA	NA	NA

Key patterns across leagues:

- Average market value varies significantly, with the Premier League showing higher valuations.
- Goals and assists are more concentrated among forwards.
- Age distribution is consistent, skewed towards younger players.

3 Model

A linear regression model was developed for each league:

$$\begin{aligned}\text{Market Value}_i = \beta_0 &+ \beta_1(\text{Age}_i) + \beta_2(\text{Goals}_i) \\ &+ \beta_3(\text{Assists}_i) + \beta_4(\text{Club Ranking}_i) \\ &+ \beta_5(\text{National Team Ranking}_i) + \beta_6(\text{Minutes Played}_i) \\ &+ \beta_7(\text{Position}_i) + \epsilon_i\end{aligned}$$

4 Results

4.1 Regression Coefficients by League

The table below presents the key regression results for each league.

Table 2: Regression Results for Premier League (England)

Variable	Estimate	Standard Error	P-Value
(Intercept)	41056662.569	4663120.476	0.000
age	-1136965.761	172607.927	0.000
goals	2992910.736	292244.210	0.000
assists	1567014.898	408471.936	0.000
club_ranking	-26533.758	3822.203	0.000
national_team_ranking	-64770.002	31992.895	0.044
minutes_played	4369.443	1034.537	0.000
positionFW	-9608394.216	2661027.166	0.000
positionFWDF	-7253141.671	4575391.947	0.114
positionFWMF	-3811156.111	2299731.089	0.098
positionMF	2018965.027	2083147.552	0.333
positionMFDF	-2041998.275	2822091.865	0.470

Table 3: Regression Results for Ligue 1 (France)

Variable	Estimate	Standard Error	P-Value
(Intercept)	22946391.141	3192522.825	0.000
age	-593454.047	128441.689	0.000
goals	2027317.425	236947.317	0.000
assists	1714839.438	429571.351	0.000
club_ranking	-15692.968	3057.507	0.000
national_team_ranking	-35781.593	18260.125	0.051
minutes_played	653.709	825.955	0.429
positionFW	-7041471.534	2084659.819	0.001

Variable	Estimate	Standard Error	P-Value
positionFWDF	-4371874.364	2938068.606	0.138
positionFWMF	-5010537.338	1642052.576	0.002
positionMF	-2937864.490	1500036.685	0.051
positionMFDF	-1905014.765	2148414.651	0.376

Table 4: Regression Results for Bundesliga (Germany)

Variable	Estimate	Standard Error	P-Value
(Intercept)	28663269.385	3916775.883	0.000
age	-848396.672	150315.678	0.000
goals	1473769.616	227410.144	0.000
assists	1827775.669	355813.328	0.000
club_ranking	-16889.707	3537.051	0.000
national_team_ranking	12408.063	35314.849	0.726
minutes_played	1351.029	941.231	0.152
positionFW	-2081053.442	2557078.632	0.416
positionFWDF	-7835403.852	3077121.174	0.011
positionFWMF	-5238301.846	1844954.498	0.005
positionMF	-1468911.257	1816906.028	0.419
positionMFDF	-2433045.981	2127195.227	0.253

Table 5: Regression Results for Serie A (Italy)

Variable	Estimate	Standard Error	P-Value
(Intercept)	26361462.597	3037789.556	0.000
age	-828777.471	112509.406	0.000
goals	2489286.770	216405.957	0.000
assists	820003.487	352805.839	0.021
club_ranking	-17977.374	2896.295	0.000
national_team_ranking	-13743.219	19749.064	0.487
minutes_played	2296.142	664.685	0.001
positionFW	-2246877.355	1693763.310	0.185
positionFWDF	-2326714.975	3227375.737	0.471
positionFWMF	-4876873.138	1510409.924	0.001
positionMF	-1910740.953	1267585.530	0.132
positionMFDF	-1933487.605	2183378.774	0.376

Table 6: Regression Results for La Liga (Spain)

Variable	Estimate	Standard Error	P-Value
(Intercept)	35886841.92	4504735.313	0.000
age	-1179214.86	160280.888	0.000
goals	1273177.50	296325.217	0.000
assists	1548163.36	501861.573	0.002
club_ranking	-33958.51	5881.304	0.000
national_team_ranking	-19577.18	36619.474	0.593
minutes_played	3508.06	1068.252	0.001
positionFW	-4474049.44	2787506.185	0.109
positionFWDF	62219.23	5844876.892	0.992
positionFWMF	2012615.10	2393250.829	0.401
positionMF	1169964.30	2029437.902	0.565
positionMFDF	6327120.85	3003669.879	0.036

4.2 Key Findings

1. **Age**: Age negatively impacts market value across all leagues, with the effect strongest in Serie A and La Liga.
2. **Goals and Assists**: These variables are the most consistent predictors of market value. Their influence is largest in the Premier League, suggesting a higher emphasis on offensive contributions.
3. **Club Ranking**: Significant in all leagues, with stronger effects observed in Ligue 1 and Serie A.
4. **Position**: Forwards generally have lower valuations compared to midfielders and defenders in Ligue 1 and La Liga.

5 Discussion

5.1 What Was Done

This study developed league-specific linear regression models to analyze the determinants of soccer players' market values.

5.2 Insights Gained

1. Offensive contributions (goals, assists) are crucial drivers of market value, but their relative importance varies by league.

2. Club ranking is consistently significant, reflecting the role of team reputation in player valuation.
3. Age is negatively associated with market value, underscoring the premium placed on youth in professional soccer.

5.3 Limitations and Future Directions

The study is limited to one season and excludes external factors like injuries, media influence, or tactical role. Future work could explore time-series data and non-linear modeling approaches.

A References