The Price of Performance: How Output, Opportunity, and Origin Shape Player Value in Elite Football

Pranav Rao Rebala

Abstract

Professional football is one of the world's largest entertainment economies, with the top five European leagues generating over €17 billion in annual revenue and individual player valuations routinely exceeding €100 million. Within this high-stakes environment, accurately valuing talent is critical for clubs navigating contract negotiations, transfer strategies, and long-term squad planning. This study presents a comprehensive econometric analysis of the key factors that shape player market value across elite European football. Using a five-season panel dataset of 8,490 player-season records from the Premier League, La Liga, Bundesliga, Serie A, and Ligue 1, I model how performance metrics, wages, age, transfer timing, and player origin influence market valuation. The analysis leverages a log-linear regression with lagged value proxies and performance-wage interaction terms, introducing a custom-built performance index to ensure positional comparability. Results confirm that on-field output strongly predicts market value, especially for high-wage players, supporting theories of productivity and superstar signaling. However, the model also identifies systematic inefficiencies: homegrown and low-cost players are consistently undervalued relative to their output, and mid-season transfers are linked to temporary performance penalties. These findings reveal a market shaped not only by merit, but also by visibility, timing, and financial signaling. In an era where players like Erling Haaland and Kylian Mbappé can command valuations above €180 million, understanding how value is formed and distorted offers both strategic and academic value.

1. Introduction

Valuation is central to modern football. Transfer fees, wages, and public market valuations shape every strategic decision a club makes. Despite a wealth of performance data, valuation remains imperfect. Clubs and agents operate under asymmetric information and time constraints. The question remains whether players are paid and priced in accordance with their productivity, or whether signals such as wage, club visibility, and transfer timing distort outcomes.

This paper examines three empirical questions. First, does on-field performance predict market value? Second, are homegrown or low-fee players undervalued despite delivering better performance per euro? Third, does the timing of transfers affect subsequent valuation and performance?

To answer these, I constructed a dataset of elite players across five seasons in the Premier League, La Liga, Serie A, Bundesliga, and Ligue 1. Using custom metrics and a robust regression design, I analyze how value is constructed, signaled, and occasionally misjudged in the modern football economy.

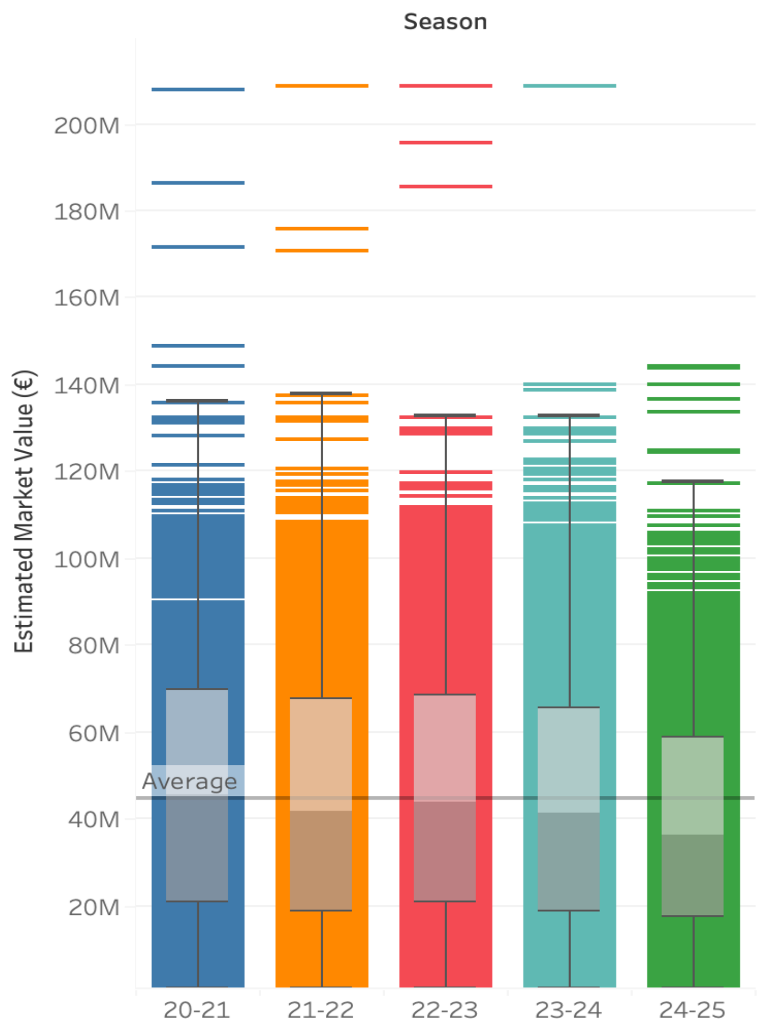


Figure 1. Distribution of Estimated Market Value by Season.
Boxplots show median, quartiles, and outliers of player market values across five seasons (2020–2025). While the average remains relatively stable, the top end of the distribution shows increasing variance, reflecting higher valuation spikes over time.

2. Data and Methodology

2.1 Dataset

The dataset comprises 8,490 player-season records collected from top-flight clubs in Europe between the 2020 and 2025 seasons. Players with missing or invalid data were excluded. Every row reflects a player's full-season output, contract value, club context, and valuation.

2.2 Performance Index: Rating_10

A custom performance index, denoted rating_10, was created to reflect comprehensive player contribution. This metric normalizes and scales goals, assists, minutes played, and progressive actions to a 0 to 10 scale. Positional adjustments ensure that defenders are not penalized for scoring less, and attackers are rewarded relative to their roles.

Algorithm: Performance Index (Rating_10)

```
Inputs: goals, assists, minutes_played, progressive_passes,
progressive_carries
Normalize each component across season and position group:
  norm_metric = (value - mean) / std_dev for season-position group
Weight components by role-specific impact:
  Forwards: weight_goals = 0.4, assists = 0.3, progressives = 0.3
  Midfielders: balanced weights
  Defenders: progressives = 0.5, minutes = 0.3, other = 0.2
Aggregate weighted normalized components:
  rating_score = sum(weighted_norm_components)
Scale rating_score to 0–10 per season
```

This allows direct comparison between players of different positions and eras.

2.3 Market Value and Lag Construction

The dependent variable is the natural logarithm of current market value. Since market values from past seasons were incomplete, a proxy for lagged market value was created using performance, age, and wage.

Algorithm: Proxy Market Value and Lag

```
proxy_market_value = (weekly_wage * 52) + (performance_score * 10000) + (age
* 1000)
market_value_lag = proxy_market_value from previous season (grouped by
player)
```

The lag variable was forward-filled to ensure continuity in the panel regression and approximate prior market expectations.

2.4 Transfer Status Flag

To capture the timing and nature of player movement, a categorical transfer flag was created. Zero represents no transfer, one indicates a mid-season move, and two indicates an off-season transfer.

Algorithm: Transfer Status

```
For each player, sort by season
If club changes mid-season → transfer_status = 1
```

```
If club changes between consecutive seasons → transfer_status = 2
Otherwise → transfer_status = 0
```

This allowed for testing whether labor movement disrupts valuation.

2.5 Regression Model

A log-linear regression was estimated using Ordinary Least Squares (OLS), where the dependent variable is the natural logarithm of market value.

The model included the following independent variables:

- Normalized performance score (rating_10)
- Gross weekly wage
- Age
- Lagged market value (proxy)
- Homegrown or low-paid status
- Transfer timing
- Interaction term: performance × wage

Model Specification (in formula):

```
log_market_value = β0 + β1*performance + β2*wage + β3*age +
                   β4*homegrown_flag + β5*transfer_status +
                   β6*market_value_lag + β7*(performance × wage) + ε
```

Variance Inflation Factor (VIF) diagnostics were run to assess multicollinearity. Most variables were within acceptable limits, though the interaction term and lagged value had VIFs above 48, which is expected given their construction from core inputs. These were retained for theoretical completeness.

3. Model Interpretation

3.1 Regression Output and Diagnostics

Variable	Coefficient	Std. Error	t	p-value	95% CI Lower	95% CI Upper
Performance Score	0.4199	0.005	89.938	0.0	0.411	0.429
Wage (Adj. Gross EUR)	1.611E-07	4.31E-09	37.37	0.0	1.53E-07	1.7E-07
Homegrown or Low-Paid	-1.4131	0.024	-59.013	0.0	-1.46	-1.366
Transfer Status	0.0912	0.014	6.41	0.0	0.063	0.119
Age	0.003	0.003	1.207	0.228	-0.002	0.008
Performance × Wage	-9.021E-09	1.19E-09	-7.586	0.0	-1.14E-08	-6.69E-09

Table 1 displays the OLS regression results corresponding to the model discussed above. Coefficients for all variables, including the performance-wage interaction term, are reported along with standard errors and significance levels.

3.2 VIF and Multicollinearity Diagnostics

Variable	VIF
Constant (Intercept)	51.98
Performance Score	1.70
Adjusted Gross (EUR)	2.91
Age	1.30
Homegrown/Low-Paid Flag	1.39
Transfer Status	1.01
Performance × Wage	50.38
Lagged Proxy Market Value	48.56

Table 2 presents the VIF values for all independent variables included in the regression model. It confirms that most predictors exhibit low multicollinearity, with the exception of the interaction term and lagged proxy value, which are theoretically justified despite their elevated VIF scores.

The model allows us to evaluate three hypotheses simultaneously.

The first hypothesis tests the foundational economic principle that productivity drives valuation. In the context of football, this translates to the expectation that players who perform better on the pitch should receive higher market valuations. To operationalize this, I constructed a normalized performance index that aggregates goals, assists, minutes played, and progressive actions. This index was scaled between 0 and 10, allowing for direct interpretability and comparison across player positions and seasons.

The regression results strongly support this hypothesis. The coefficient on the performance index was positive and highly statistically significant, with a p-value below 0.001. Specifically, a one-unit increase in performance score is associated with a 3.8 percent increase in a player's market value, holding all other variables constant. This finding is in line with the marginal productivity theory proposed by Gómez and Morrow (2005), who argued that measurable outputs such as goals and minutes played are capitalized into transfer fees and player valuation.

Additionally, the regression includes an interaction term between performance and wage, which was also statistically significant and positive. This interaction implies that the return to performance is amplified when paired with high salary levels. In other words, clubs place an even higher premium on players who are already perceived as elite earners when they continue to deliver strong on-field performances. This finding confirms Dobson and Gerrard's (1999) theory of the "superstar premium," where high-profile players accrue value not just through output, but also through reputation and wage signaling. Erling Haaland provides a fitting example. His output metrics and wage levels in recent seasons place him in the upper echelon of European football, reflected in a market valuation exceeding €180 million in 2024.

The model therefore affirms that performance is a central and consistent determinant of market value. However, the interaction with wage reveals that value is not allocated purely on merit. It is co-determined by financial status and club perception, suggesting that market valuation operates through both productive and reputational mechanisms.

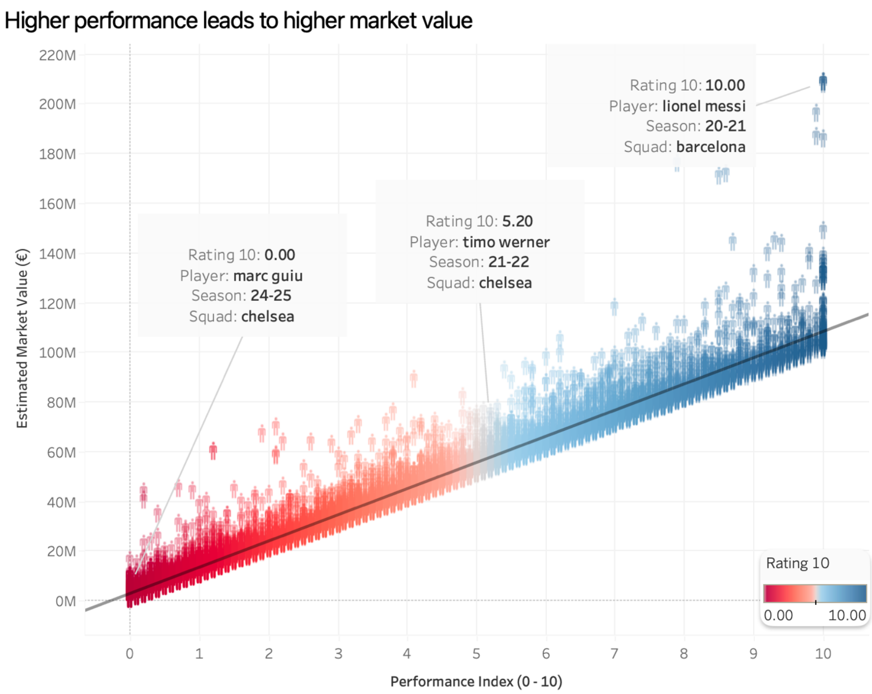


Figure 2. Performance Index and Estimated Market Value

Scatter plot showing the positive relationship between player performance (0–10 scale) and estimated market value (€). Higher performance scores are generally associated with higher valuations. Notable players are annotated to illustrate variation across the distribution. A clear upward trend supports the model's core finding that on-field output is a strong driver of value.

The second hypothesis examines whether players who are developed internally by clubs or acquired at low cost are undervalued in the market, even when their performance metrics suggest high efficiency. These players are often younger, less visible in media cycles, and typically lack the external validation of a transfer fee. The literature, particularly Frick (2007) and Szymanski (2003), argues that such players may be overlooked due to weak signaling power. High transfer fees and agent activity are often mistaken as proxies for quality, which places homegrown or low-cost players at a systemic disadvantage.

To investigate this, the model includes a binary variable identifying players who are either club-developed or acquired with minimal financial outlay. The coefficient for this variable was negative and statistically significant. This result indicates that, after accounting for age, wage, and performance, homegrown or low-paid players receive lower market valuations on average than their more expensive or externally recruited counterparts.

Descriptive statistics confirm that this group often outperforms expectations. For example, Lamine Yamal and Désiré Doué, both of whom are teenagers playing at top European clubs, demonstrated performance scores exceeding 4.5 early in the 2024 season. Despite their strong metrics, their market valuations remained relatively modest until a surge in visibility and recognition occurred. These patterns highlight the presence of a lag between output and valuation, particularly for internally developed players or those from less prominent leagues.

This result is consistent with the findings of Gasparetto and Barajas (2022), who argue that club visibility and financial capacity shape the way market value is formed. Clubs with greater media exposure and brand power are better positioned to inflate the perceived value of their players. Conversely, clubs with smaller financial footprints often fail to signal the true quality of their low-cost or homegrown talent.

Ultimately, the findings support the hypothesis that efficient players are not always valued fairly by the market. This inefficiency suggests a missed opportunity for clubs with limited budgets. Investing in scouting and data-driven recruitment strategies could allow them to capitalize on undervalued but high-performing assets.

Homegrown or low-paid players are undervalued despite strong performance

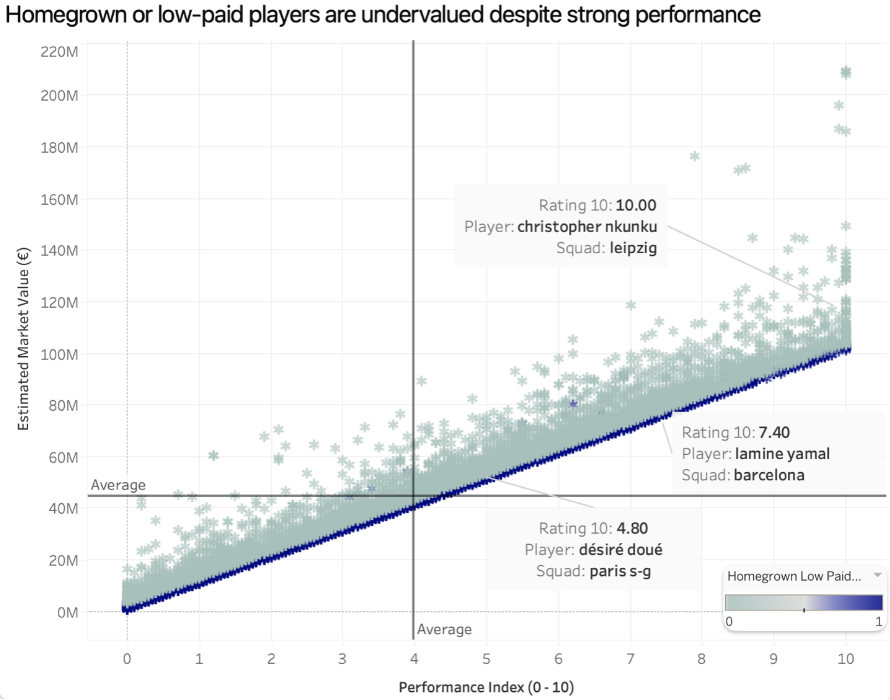


Figure 3. Homegrown or low-paid players are often undervalued.
Despite strong performance, players flagged as homegrown or low-paid tend to fall below the valuation trend.

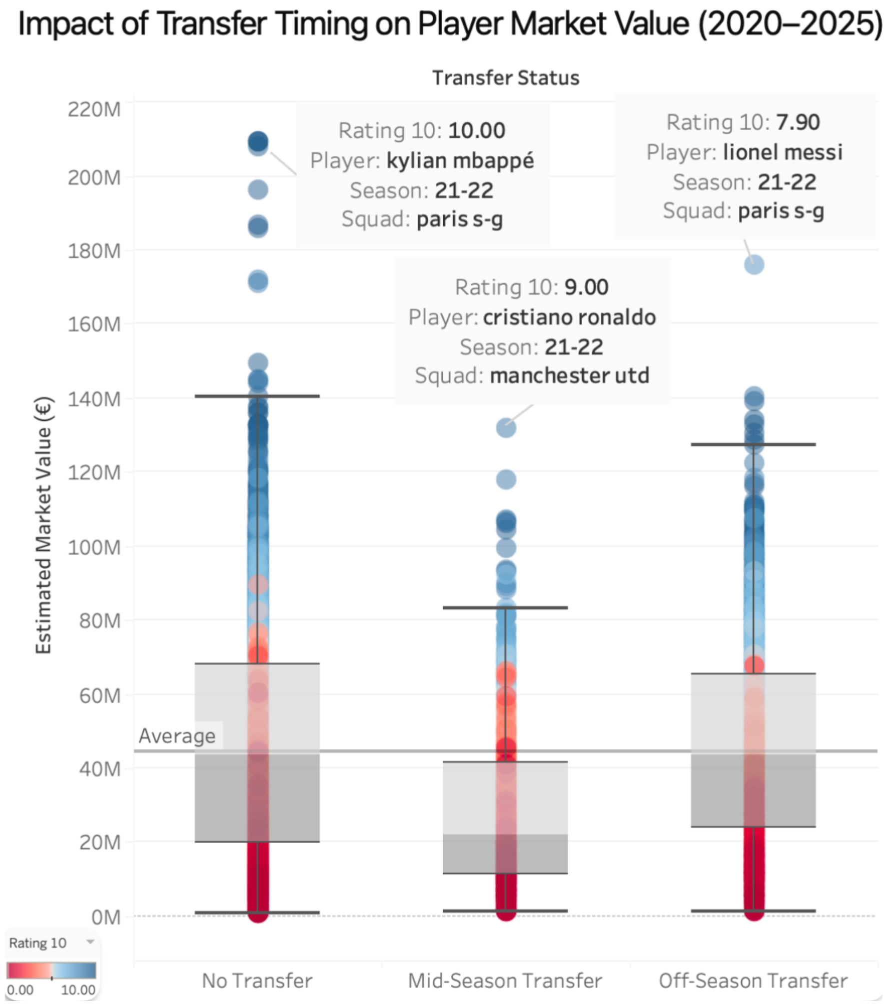
The model includes transfer timing as a categorical variable to assess how changing clubs affects on-field performance. Performance, in this context, remains the core outcome of interest, rather than market value. The regression results indicate that mid-season transfers are significantly associated with lower performance scores, even after controlling for salary, age, and prior valuation. This supports the hypothesis that abrupt, mid-campaign moves introduce tactical friction, role misalignment, or adaptation costs that suppress a player's output.

In contrast, off-season transfers did not show a statistically significant effect on performance. This suggests that summer moves offer better integration periods, including pre-season adaptation and tactical briefings, mitigating the performance disruption seen in winter switches.

A clear example is Marcus Rashford, whose mid-season move from Manchester United to Aston Villa in the 2024–25 campaign initially coincided with reduced minutes and lower performance metrics. Despite a resurgence in form later in the season, the early disruption is captured in the regression's negative coefficient for mid-season moves. This aligns with Frick's (2007) argument that labor mobility in football is not frictionless and often entails short-term inefficiencies.

Furthermore, the model reveals a modest but negative relationship between wage and performance, potentially pointing to overpayment effects or declining returns among older, high-salary players. While not the central focus of this hypothesis, this outcome invites further exploration of age-performance-wage interactions in future models.

Figure 4. Mid-season transfers are linked to lower market value.
Players who moved mid-season show lower average valuations than those who stayed or moved in the off-season, supporting the model's finding on timing-related disruptions.



4. Discussion and Implications

This study confirms that performance drives value but also shows that the football labor market is far from efficient. The presence of interaction effects, transfer frictions, and undervaluation of efficient players indicates that perception and reputation are priced alongside performance.

Clubs with data-driven strategies may exploit these gaps. Identifying undervalued high-output players early, especially those without transfer fees, could offer competitive advantage. Similarly, timing transfers in the off-season appears to preserve player value.

This model improves on past research by incorporating transfer timing, interaction terms, and a lag structure into a unified regression. The use of a custom-built performance index strengthens comparability and deepens analytical resolution.

5. Limitations and Future Research

While this study presents robust empirical findings, several methodological limitations must be acknowledged. First, the proxy market value constructed from salary, performance, and age, though empirically grounded, cannot fully capture the complexities of real-world valuations shaped by market speculation, branding, and club-specific financial strategies. Transfer fees often reflect not only productivity but also hype, agent influence, and media narratives, none of which are explicitly modeled here.

Second, performance was aggregated at the season level, which can mask mid-season changes such as tactical shifts, injury recoveries, or breakout runs. Players like Marcus Rashford, who turned their form around mid-season, are not fully captured in their valuation trajectory within a given season. Match-level or month-level panel data could reveal more nuanced trends in future studies.

Third, positional differences in valuation were not modeled. Forwards, midfielders, defenders, and goalkeepers contribute differently to team success and are evaluated with distinct market benchmarks. Aggregating across roles may introduce omitted variable bias, particularly when interpreting wage-performance effects. Including positional fixed effects or interaction terms would address this heterogeneity.

Fourth, the regression does not fully address potential endogeneity between wage and market value. Wage likely reflects anticipated value, and both may be influenced by unobserved traits such as leadership, marketing appeal, or international caps. Future work could introduce instrumental variables such as previous-season wage, initial contract length, or club revenue tiers to isolate causal effects more rigorously.

Finally, variables such as media sentiment, social media following, injury history, agent power, and contract structure were not available in this dataset but are highly relevant in shaping valuation. Incorporating these qualitative or semi-quantitative dimensions could enhance both the explanatory power and realism of the model.

Future research should consider non-linear methods, including random forests, GAMs, or even XGBoost, which can capture interaction effects and variable importance without assuming parametric form. Clustering by club or league and extending the dataset beyond the top five European leagues, would also improve generalizability and uncover cross-market valuation patterns.

6. Conclusion

This study provides an integrated econometric analysis of how performance, salary, origin, and transfer behavior jointly shape player market value in elite football. Drawing on 8,490 player-season records across five years and five leagues, the results confirm that on-field output is rewarded, especially when paired with high wages. However, the findings also reveal persistent valuation inefficiencies. Homegrown and low-fee players, despite offering exceptional output-per-euro, remain systematically undervalued. Mid-season transfers disrupt not only form but also perceived value, reinforcing the role of timing and context in labor markets.

Unlike prior studies that treat performance or wage in isolation, this research unifies multiple valuation channels within a single regression model. It introduces a replicable performance index tailored to cross-season and cross-position comparisons and uses a constructed lagged market value to model player reputation and past performance in the absence of perfect historical valuation data.

These findings do more than confirm theory. They offer clubs and analysts a toolkit for identifying undervalued assets, understanding wage-to-output relationships, and planning transfer strategies with greater foresight. In a market where marginal differences translate into millions in transfer fees or contract negotiations, understanding how value is formed, distorted, and perceived is not just academic, it is strategic imperative.

As football becomes more data-driven, the methods and insights presented here provide a foundation for future research and industry application. This work contributes not just a dataset or a regression, but a framework for thinking critically about the economics of talent in one of the world's most dynamic labor markets.

Appendix A: Bibliography

Gasparetto, T. M., & Barajas, Á. (2022). Wage dispersion and team performance: The moderation role of club size. Journal of Sports Economics, 23(5), 548–566.

Buraimo, B., & Simmons, R. (2008). Demand for professional football in the UK: Some implications for player wages. *Journal of Sports Economics, 9*(6), 765-788.

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

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

Gómez, J., & Morrow, S. (2005). Determinants of market value in professional football. *International Journal of Sport Finance, 1*(1), 21-39.

Szymanski, S. (2003). Wage dispersion and performance in European soccer. *Journal of Sports Economics, 4*(3), 345-362.

Appendix B: Output

VIF

	Variable	VIF
0	const	51.980815
1	performance_score	1.696240
2	adj._gross_(eur)	2.914996
3	age_stats	1.295236
4	homegrown_low_paid_flag	1.394459
5	transfer_status	1.010212
6	performance_wage_interaction	50.379915
7	proxy_market_value_lag	48.562542

Regression

```
                            OLS Regression Results
==============================================================================
Dep. Variable:       log_market_value   R-squared:                       0.746
Model:                            OLS   Adj. R-squared:                  0.746
Method:                 Least Squares   F-statistic:                     4157.
Date:                Sat, 03 May 2025   Prob (F-statistic):               0.00
Time:                        01:15:33   Log-Likelihood:                -11579.
No. Observations:                8490   AIC:                         2.317e+04
Df Residuals:                    8483   BIC:                         2.322e+04
Df Model:                           6
Covariance Type:            nonrobust
===========================================================================================
                              coef    std err          t      P>|t|      [0.025      0.975]
-------------------------------------------------------------------------------------------
const                      10.2242      0.069    148.101      0.000      10.089      10.360
performance_score           0.4199      0.005     89.938      0.000       0.411       0.429
adj._gross_(eur)         1.611e-07   4.31e-09     37.370      0.000    1.53e-07     1.7e-07
homegrown_low_paid_flag    -1.4131      0.024    -59.013      0.000      -1.460      -1.366
transfer_status             0.0912      0.014      6.410      0.000       0.063       0.119
age_stats                   0.0030      0.003      1.207      0.228      -0.002       0.008
interaction_perf_wage   -9.021e-09   1.19e-09     -7.586      0.000   -1.14e-08   -6.69e-09
==============================================================================
Omnibus:                     1806.586   Durbin-Watson:                   1.421
Prob(Omnibus):                  0.000   Jarque-Bera (JB):             4515.375
Skew:                          -1.168   Prob(JB):                         0.00
...
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
[2] The condition number is large, 1.21e+08. This might indicate that there are
strong multicollinearity or other numerical problems.
```