

Optimal Replacement of Harry Kane: An Empirical Model of Daniel Levy

Satyam Anand Axel Anderson Jeremy Rosen

Georgetown University

February 14, 2025

TransferMarkt does not publish data for female players.
This presentation consequently only concerns men's football.
We will be using "he/him/his" when referring to players.

Table of Contents

- 1 Introduction to Football
- 2 Research Questions
- 3 Model for Contests
- 4 Estimation Results
- 5 Conclusion

Introduction to Football

- *"Il calcio è la cosa più importante delle cose meno importanti"*
Football is the most important of the least important things
– Arrigo Sacchi, *Italy manager ('91 – '96)*
- Football clubs in Europe are worth billions of euros with players' transfer fees often exceeding tens of millions
- Several proposed reforms to address concerns about long-term financial sustainability and workers' rights
- Semi-transparent setting for labour mobility with heterogeneous firms and productivity

Football Clubs as Firms

- Clubs across Europe operate in domestic competitions (leagues) structured in a tiered pyramid format (divisions) with around 20 participants in each time period (season) between August and May
 - Top- and bottom-placed performers (generally 3 each) are promoted and relegated into higher and lower divisions respectively
 - Major revenues gaps exist across divisions and countries, more so in England, France, Germany, Italy and Spain (Big 5)
 - Clubs may only register 25 players aged 21+ in general

Example

Leicester City finished first in the English 2nd Division in 2023-24 and were consequently promoted to the English 1st Division for 2024-25.

Competition in Football

- Open League (promotion/relegation) creates pressure, no explicit mechanism to guarantee competitive balance (e.g. draft system in North America, aerodynamics in Formula One)
- Top ranked clubs (generally up to 6th place in the Big 5) qualify for tournaments hosted by UEFA, the Europe-wide governing body
 - Earnings from UEFA are comparable to domestic equivalent, may vary by structure of broadcasting agreements
- Elite clubs must compete domestically and internationally, which increases demand for high quality players

▶ Revenues

Research Questions

The Curious Case of Harry Kane



Harry Kane started his career at Tottenham Hotspur in 2013.

He is the second highest goalscorer in the history of the Premier League and the captain of England since 2018.

However, Spurs have failed to win any tournaments since 2008. Frustrated, Kane submitted a transfer request in May 2021 to leave Spurs.

(Image Source: WikiCommons)

Daniel Levy's Dilemma

- Historically an upper mid-table team, Spurs with Kane as striker began to challenge for (but not win) the Premier League
 - Spurs and Man City now included in the “Big Six” along side Man Utd, Liverpool, Arsenal, and Chelsea
- Owner and chairman Daniel Levy must decide if Kane should be sold, and potentially find a suitable replacement
 - ① How much does Kane contribute to Spurs' on-field success?
 - ② How much will Kane grow as a player in the foreseeable future?
 - ③ Which players would be adequate replacements for Kane?
 - ④ How much should be paid to acquire any prospective signing?

- Is the transfer market in European football efficient?
 - ① How do clubs determine each player's value for their team?
Estimate production functions and calculate "marginal points"
 - ② How do player performances influence valuations?
Estimate valuations based on player statistics and other attributes
 - ③ How do clubs search for prospective players?
Identify determinants of transfer activity in the Big 5 using previously estimated parameters

- Even more broadly, how do firms...
 - 1 Determine individual workers' values in a group setting?
 - 2 Search for prospective acquisition targets?
 - 3 Negotiate transaction fees and wages for workers?

- 32K+ matches, 16K+ players, 200+ clubs since 2006-07
 - Analysis estimates heterogeneous production functions for clubs with at least one season in the Big 5 since 2006-07
- Merged with betting odds data from two major UK bookmakers to obtain continuous outcome variable, “expected points”
- Additional available data includes info for 25K+ players from 17 leagues as well as financial data for Big 5 clubs since 2017-18

Data Source: TransferMarkt

- German sports website that provides “crowd-sourced” valuations for football players as well as transfers, rumors and other information
- “Market Value” relies on “wisdom of crowds”; no strict definition
 - Viewed as transfer fee required for player with 2-3 years left on contract

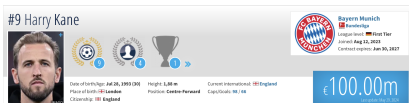


Figure 1: Player Profile

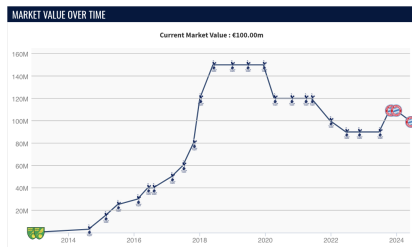


Figure 2: Market Value

Model for Contests

- Is the transfer market in European football efficient?
 - ① **How do clubs determine each player's value for their team?**
Estimate production functions and calculate “marginal points”
 - ② How do player performances influence valuations?
Estimate valuations based on player statistics and other attributes
 - ③ How do clubs search for prospective players?
Identify determinants of transfer activity in the Big 5 using previously estimated parameters

Model for Contests: Budgets and Costs

- For now, assume that clubs operate within a static framework
 - Clubs attempt to maximize expected points across the season,
 - Clubs view themselves in isolation; cannot influence others' points (except for the fixtures pertaining to the specific pair of teams)
 - Endowments e_t^A are based on league position in the previous year $t - 1$
 - Each player must be paid wages $w_g(z_{p,t}^A)$ based on their grouping g
 - Additionally, the transfer fees to acquire players must be linearly amortized across their remaining contract, $\phi(z_{p,t}^{A,\tau})$
- **Current Objective:** Model an individual fixture where both teams select players and attempt to maximize their expected payoff.

Fixtures and Line-Ups

- Define a **fixture** i between Team A and Opponent B at venue $v \in \{A, B\}$ in season t , i.e. $i = \langle A, B, v, t \rangle$
 - For example, $i = \langle \text{Tottenham, Arsenal, Tottenham, 2023-24} \rangle$
- Each club $k \in \{A, B\}$ selects its fixture-specific starting line-up $z_i^k = \{z_{i,p}^k\}_{p=1}^P$ where $z_{i,p}^k$ denotes all players' valuations
 - P denotes the number of players (eleven starting, up to nine benched)
- Players can be assigned to position groups $g \in \{G, D, M, F\}$, which correspond to goalkeepers, defenders, midfielders and forwards
- Teams select their line-ups to maximize their expected points y_i^k

Expected Points

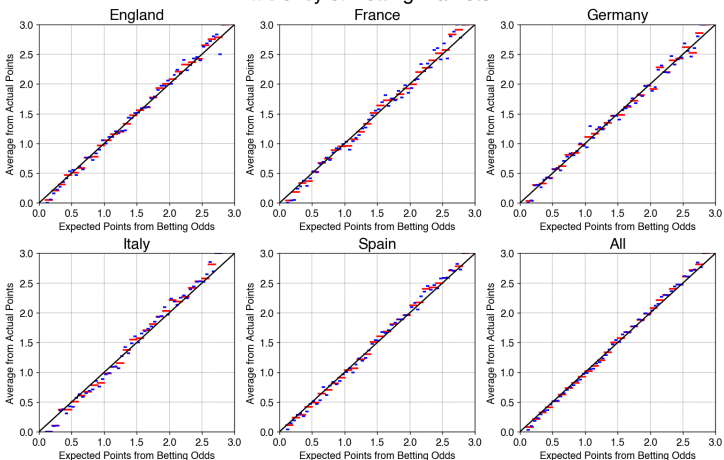
- Clubs are assumed to maximize their expected points y_i^k by selecting optimal line-ups and strategies to compete against their opponents
- Expected points from a fixture i can be inferred from the odds provided by betting companies $b \in \{\text{Bet365}, \text{BetVictor}\}$
 - Sum of implied probabilities exceeds 1 (known as bookmaker's edge); re-scaled probabilities may be used to compute expected points instead

$$y_{i,b}^A = \mathbb{P}_{i,b}^W \cdot \pi_W + \mathbb{P}_{i,b}^L \cdot \pi_L + \mathbb{P}_{i,b}^D \cdot \pi_D$$
$$y_{i,b}^{A*} = \frac{1}{\sum \mathbb{P}_{i,b}} \cdot \left[\mathbb{P}_{i,b}^W \cdot \pi_W + \mathbb{P}_{i,b}^L \cdot \pi_L + \mathbb{P}_{i,b}^D \cdot \pi_D \right] \quad (1)$$

- Average expected points across both bookmakers were used

Efficiency of Betting Markets

Efficiency of Betting Markets



- Markets align with results; tail behavior due to longshot bias

Position Groups and Direct Effects

- Players' valuations $z_{i,p}^k$ can be aggregated into position groups $x_{i,g}^k$ using the generalized mean $f(\dots)$ parameterized by α_g

$$x_{i,g}^k = f(z_{i,g}^k; \alpha_g) = \left[\frac{1}{P_{i,g}^k} \cdot \sum_{z_{i,p}^k \in z_{i,g}^k} (z_{i,p}^k)^{\alpha_g} \right]^{\frac{1}{\alpha_g}} \quad (2)$$

- $z_{i,g}^k \equiv \{z_{i,p}^k \in z_i^k : g(z_{i,p}^k) = g\} \subset z_i^k$ where $g(\dots)$ denotes an indicator-like mapping to the various position groups g

Example

Romero (D), Porro (D), Werner (F), Son (F) will be aggregated into their position groups (D and F). The grouping function may vary by position group via the aggregation parameter, i.e. $\alpha_g \neq \alpha_{g'}$.

Interaction Terms

- The expected points for Team A can be expressed directly as a linear function of both clubs' position group valuations

$$y_i^{A*} = \sum_{k \in \{A, B\}} \left[\sum_g \beta_g^k \cdot x_{i,g}^k \right] + \varepsilon_i^A \quad (3.a)$$

- Efficiencies in aggregation across position groups **within** club k can be modeled by including interaction terms

$$y_i^{A*} = \sum_k \left[\sum_g \beta_g^k \cdot x_{i,g}^k + \sum_{g \neq g'} \gamma_{gg'} \cdot x_{i,g}^k \cdot x_{i,g'}^k \right] + \varepsilon_i^A \quad (3.b)$$

Example

β_D, β_F denote the direct contribution toward expected points whereas the interaction term γ_{DF} denotes the synergy between defenders and forwards within a club (e.g. fullbacks contributing by crossing to forwards).

Opposition-Based Spillovers

- Competition between overlapping positions groups across opponents can similarly be modeled by including additional interaction-like terms

$$\begin{aligned} y_i^{A*} &= \sum_k \left[\sum_g \beta_g^k \cdot x_{i,g}^k + \sum_{g \neq g'} \gamma_{gg'} \cdot x_{i,g}^k \cdot x_{i,g'}^k \right] \\ &+ \delta_{GF} \cdot \frac{x_{i,G}^A}{x_{i,F}^B} + \delta_{DF} \cdot \frac{x_{i,D}^A}{x_{i,F}^B} + \delta_{MM} \cdot \frac{x_{i,M}^A}{x_{i,M}^B} \\ &+ \delta_{FD} \cdot \frac{x_{i,F}^A}{x_{i,D}^B} + \delta_{FG} \cdot \frac{x_{i,F}^A}{x_{i,G}^B} + \varepsilon_i^A \end{aligned} \quad (3.c)$$

Example

δ_{GF} denotes the direct effect of goalkeepers impeding opposing forwards (e.g. a goalkeeper's shot-stopping ability diminishes forwards' goal-scoring), δ_{MM} denotes the direct effect of midfielders battling against each other (e.g. defensive midfielders intercepting attacking midfielders) and so on.

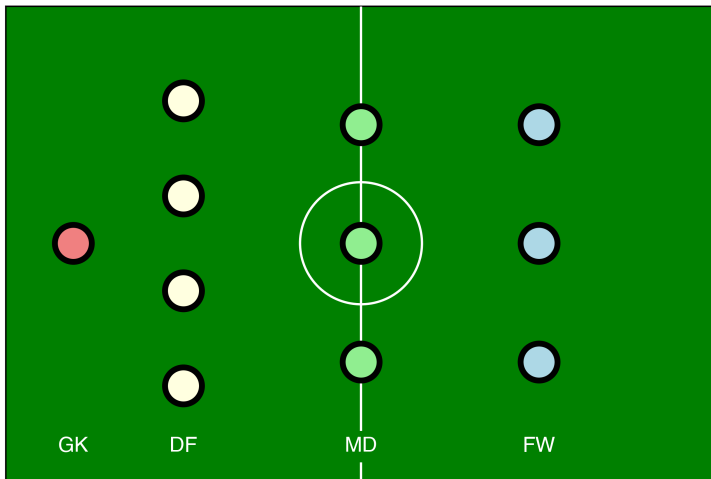
Formations, Managers and Home Advantage

- Teams are organized in “formations” that are titled based on the number of players assigned to each position grouping “D-M-F”
 - “4-3-3” (40%+) and “4-4-2” (25%+) formations are commonly used
- “Managers” are responsible for determining team tactics; fixed effects were constructed for managers with at least 50 fixtures in a league
 - Managers are highly co-linear with clubs; for informativeness, manager fixed effects are used instead of club-specific fixed effects
- Home-field advantage is accounted for using a dummy variable
- Season fixed effects are also included, but generally not significant

▶ Summary Statistics

"4-3-3" Formation

4-3-3 Formation



Model Summary

- The model has team-**specific** parameters for the direct effect of each position group (β) and cross-group synergy (γ)
- Additional team-**agnostic** parameters have been included for the positional aggregations (α) and opposition-based spillovers (δ)

$$\begin{aligned}
 y_i^{A*} = & \sum_k \left[\underbrace{\sum_g \beta_g^k \cdot x_{i,g}^k}_{\text{Direct Effect}} + \underbrace{\sum_{g \neq g'} \gamma_{gg'} \cdot x_{i,g}^k \cdot x_{i,g'}^k}_{\text{Cross-Group Synergy}} \right] \\
 + & \delta_{GF} \cdot \frac{x_{i,G}^A}{x_{i,F}^B} + \delta_{DF} \cdot \frac{x_{i,D}^A}{x_{i,F}^B} + \delta_{MM} \cdot \frac{x_{i,M}^A}{x_{i,M}^B} \\
 + & \underbrace{\delta_{FD} \cdot \frac{x_{i,F}^A}{x_{i,D}^B} + \delta_{FG} \cdot \frac{x_{i,F}^A}{x_{i,G}^B}}_{\text{Opposing Spillovers}} \\
 + & \underbrace{\xi_f^A + \xi_f^B}_{\text{Formation FEs}} + \underbrace{\xi_m^A + \xi_m^B}_{\text{Manager FEs}} + \underbrace{\xi_h}_{\text{Home FE}} + \underbrace{\xi_t}_{\text{Season FE}} + \varepsilon_i^A
 \end{aligned} \tag{3}$$

- Estimate with nonlinear least squares, as opposed to OLS, due to α

Model for Contests: Optimization Problems

- For each fixture i , Team A selects a roster z_i^A of players $z_{i,p}^A$ from its squad z_t^A that maximizes its expected points from betting odds y_i^{A*}

$$\max_{z_i^A} \mathbb{E}[y_i^{A*}] \quad \text{s.t.} \quad z_{i,p}^A \in z_t^A \quad \forall z_{i,p}^A \in z_i^A \quad (\text{i.e., } z_i^A \subset z_t^A)$$

- Across the full season t , Team A selects a squad z_t^A to maximize its expected total points w.r.t its endowment e_t^A player-specific wages $w_g(z_{t,p}^A)$ and transfer fee amortization $\phi(z_{t,p}^{A,\tau})$

$$\max_{z_t^A} \sum_i \mathbb{E}[y_i^{A*}] \quad \text{s.t.} \quad \sum_{z_{t,p}^A \in z_t^A} [w_g(z_{t,p}^A) + a(z_{t,p}^A)] \leq e_t^A$$

Estimation Results

Homogeneous Estimates: α and AME

Table 1: Homogeneous Estimates with Expected Points and Season Fixed Effects

	England	France	Germany	Italy	Spain	All
α_D	.9504*** (.0976)	1.1408*** (.1125)	.7740*** (.0885)	.4730*** (.0612)	.6998*** (.0782)	.7637*** (.0364)
α_M	1.2750*** (.0988)	1.0168*** (.0916)	.6956*** (.0793)	.6739*** (.0552)	.8196*** (.1027)	.8533*** (.0370)
α_F	.6035*** (.0974)	1.1789*** (.1277)	.2833*** (.0651)	1.2460*** (.1524)	.9914*** (.0740)	.7011*** (.0399)
AME(Value _G)	.0230*** (.0021)	.0198*** (.0019)	.0081*** (.0023)	.0080*** (.0017)	.0145*** (.0019)	.0141*** (.0008)
AME(Value _D)	.1313*** (.0054)	.1000*** (.0045)	.1176*** (.0050)	.1070*** (.0040)	.0920*** (.0041)	.1076*** (.0020)
AME(Value _M)	.1430*** (.0051)	.1133*** (.0042)	.1224*** (.0046)	.1432*** (.0042)	.0782*** (.0036)	.1146*** (.0019)
AME(Value _F)	.0876*** (.0044)	.0929*** (.0034)	.1054*** (.0039)	.0808*** (.0032)	.1058*** (.0034)	.0917*** (.0016)
League FE	No	No	No	No	No	Yes
Observations	13,566	13,124	11,004	13,638	13,382	64,714
R ²	.8846	.8663	.8712	.8747	.8934	.8778
Adjusted R ²	.8834	.8649	.8695	.8732	.8921	.8769

*p<0.10, **p<0.05, ***p<0.01. Robust standard errors are in parentheses.

Homogeneous Estimates: β and γ

	England	France	Germany	Italy	Spain	All
Value _G	-.1693*** (.0575)	-.0559 (.0489)	-.2867*** (.0519)	.0521 (.0356)	.0233 (.0420)	-.0242 (.0198)
Value _D	-.2930*** (.1128)	.0142 (.0938)	-.3740*** (.1037)	-.3636*** (.0719)	.0863 (.0816)	-.3208*** (.0390)
Value _M	-.8588*** (.1123)	-.4879*** (.0891)	-.4646*** (.1071)	-.4999*** (.0762)	-.1520** (.0731)	-.4094*** (.0380)
Value _F	-.1138 (.1034)	-.6073*** (.0804)	-.6085*** (.0948)	-.2880*** (.0640)	-.9729*** (.0695)	-.4844*** (.0336)
Value _G × Value _D	.0007 (.0053)	.0009 (.0043)	.0019 (.0048)	.0100*** (.0030)	.0027 (.0037)	.0042** (.0017)
Value _G × Value _M	-.0013 (.0048)	.0024 (.0039)	.0186*** (.0045)	-.0053* (.0031)	.0029 (.0033)	.0004 (.0017)
Value _G × Value _F	.0150*** (.0055)	.0005 (.0033)	-.0025 (.0042)	-.0050* (.0027)	-.0027 (.0033)	.0001 (.0016)
Value _D × Value _M	.0478*** (.0071)	.0108* (.0057)	.0046 (.0057)	.0239*** (.0037)	-.0317*** (.0050)	.0146*** (.0022)
Value _D × Value _F	-.0216*** (.0065)	.0066 (.0050)	.0263*** (.0057)	-.0007 (.0040)	.0295*** (.0042)	.0115*** (.0021)
Value _M × Value _F	.0159*** (.0061)	.0280*** (.0047)	.0157*** (.0057)	.0208*** (.0041)	.0379*** (.0042)	.0173*** (.0020)

*p<0.10, **p<0.05, ***p<0.01. Robust standard errors are in parentheses.

Homogeneous Estimates: α , AMEs, and γ

- Models generally have good fit across all leagues ($R_{\text{adj}}^2 \geq 0.85$)
- The generalized mean parameter α is significant across all league and position combinations and approximately 1
 - Approximately between a geometric or arithmetic mean ($\alpha \in [0, 1]$)
- Average marginal effects are positive and significant across all league and position group combinations
 - AMEs for Value_D and Value_M are higher than for Value_G and Value_F for all leagues except Spain, where Value_F is highest
 - AMEs for Value_G are substantially lower; reflects goalkeepers' limited ability to contribute toward wins (as opposed to draws)
- Interaction term between midfielders and forwards γ_{MF} is positive and significant across all league and position combinations

Homogeneous Estimates: δ and Formations

	England	France	Germany	Italy	Spain	All
Value _G /Value _{-F}	-.6717 (.5535)	.2755 (.4402)	.2625 (.4953)	-.5261 (.3635)	-.8347** (.3959)	-.5192*** (.1886)
Value _D /Value _{-F}	-.1070 (.7911)	-2.9265*** (.6361)	-.2681 (.6743)	-.5970 (.4911)	-.1357 (.5725)	-.6602** (.2762)
Value _M /Value _{-M}	.1049 (.8272)	-.3497 (.6353)	-.1149 (.7339)	.5109 (.5047)	1.1238** (.5134)	.3635 (.2691)
Value _F /Value _{-D}	1.6506** (.8148)	3.0070*** (.6635)	2.1454*** (.7057)	1.2348*** (.4572)	1.6385*** (.5060)	2.0950*** (.2545)
Value _F /Value _{-G}	-.8056*** (.2734)	-.4700** (.2218)	-.5315* (.2788)	.7333*** (.1578)	-.3839* (.2244)	-.1002 (.0932)
4-4-2 Formation	-.0213*** (.0045)	-.0149*** (.0044)	-.0188*** (.0049)	-.0132** (.0053)	-.0078 (.0049)	-.0154*** (.0021)
4-2-4 Formation	.0022 (.0073)	.0046 (.0059)	.0149** (.0060)	.0138 (.0115)	.0211*** (.0050)	.0154*** (.0029)
3-5-2 Formation	.0061 (.0099)	.0011 (.0092)	-.0010 (.0086)	.0064 (.0057)	.0051 (.0116)	.0027 (.0035)
4-5-1 Formation	-.0524*** (.0077)	-.0248*** (.0062)	-.0453*** (.0085)	-.0122 (.0086)	-.0291*** (.0087)	-.0317*** (.0034)
3-6-1 or 3-4-3 Formation	.0354*** (.0105)	.0099 (.0109)	-.0091 (.0095)	-.0154* (.0086)	-.0074 (.0112)	-.0072* (.0043)
5-3-2 or 5-2-3 Formation	-.0595*** (.0085)	-.0298*** (.0081)	-.0437*** (.0123)	-.0119 (.0109)	-.0038 (.0068)	-.0259*** (.0039)
Other Formation	-.0417** (.0171)	-.0265 (.0185)	-.0324 (.0299)	-.0014 (.0298)	.0014 (.0117)	-.0194** (.0076)

*p<0.10, **p<0.05, ***p<0.01. Robust standard errors are in parentheses.

Homogeneous Estimates: Managers and Home FE

	England	France	Germany	Italy	Spain	All
Manager ₁	Pep Guardiola .3908*** (.0159)	Franck Haise .2045*** (.0173)	Urs Fischer .2390*** (.0168)	Massimiliano Allegri .1701*** (.0107)	Pep Guardiola .3261*** (.0174)	Pep Guardiola .3829*** (.0101)
Manager ₂	Alex Ferguson .3583*** (.0124)	Pascal Gastien .1497*** (.0230)	Pep Guardiola .2340*** (.0238)	Rudi Garcia .1605*** (.0200)	Jose Mourinho .2856*** (.0191)	Alex Ferguson .3528*** (.0104)
Manager ₃	Jurgen Klopp .3095*** (.0154)	Rene Girard .1286*** (.0120)	Jupp Heynckes .2051*** (.0159)	Vincenzo Italiano .1528*** (.0153)	Luis Enrique .2252*** (.0212)	Jurgen Klopp .2761*** (.0098)
Manager _{M-2}	Steve Bruce -.0820*** (.0105)	Elie Baup -.0446*** (.0141)	Michael Frontzeck -.0684*** (.0126)	Roberto De Zerbi -.0823*** (.0163)	Quique Sanchez Flores -.0403*** (.0104)	Steve Bruce -.0769*** (.0100)
Manager _{M-1}	Gareth Southgate -.0846*** (.0170)	Mehmed Bazdarevic -.0926*** (.0165)	Oliver Glasner -.0725*** (.0202)	Alberto Malesani -.0897*** (.0171)	Fran Escriba -.0713*** (.0131)	Sandro Schwarz -.0796*** (.0144)
Manager _M	Dean Smith -.0892*** (.0208)	Patrick Vieira -.1641*** (.0175)	Heiko Herrlich -.0908*** (.0196)	Alessio Dionisi -.1077*** (.0182)	Lucas Alcaraz -.1022*** (.0132)	Mehmed Bazdarevic -.0979*** (.0163)
Home FE	.4112*** (.0035)	.4720*** (.0032)	.4151*** (.0037)	.4440*** (.0034)	.4772*** (.0032)	.4448*** (.0015)

*p<0.10, **p<0.05, ***p<0.01. Robust standard errors are in parentheses.

▶ Actual Points

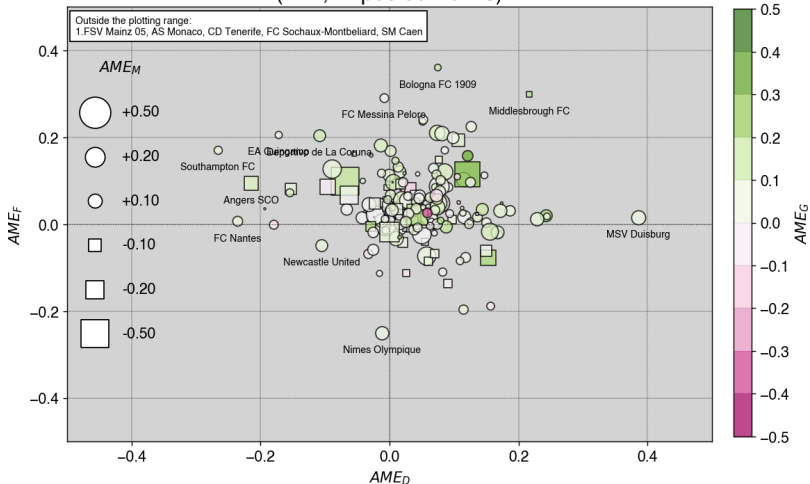
▶ Heterogeneous Estimates

Homogeneous Estimates: δ , Formations, Managers

- Interaction between own forwards and opposing defenders $\delta_{F,-D}$ is positive and significant across all leagues
 - Suggests forwards are most important against weaker defenders
- Defensive formations (e.g. “4-4-2”) perform worse than the “4-3-3”;
Offensive formations (e.g. “4-2-4”) perform better in certain leagues
- Elite managers have a significant impact on team performance
- Home field advantage is significant across all leagues and worth approximately 0.40 points (almost halfway toward a draw)

Heterogeneous Estimates: AMEs (Expected Points)

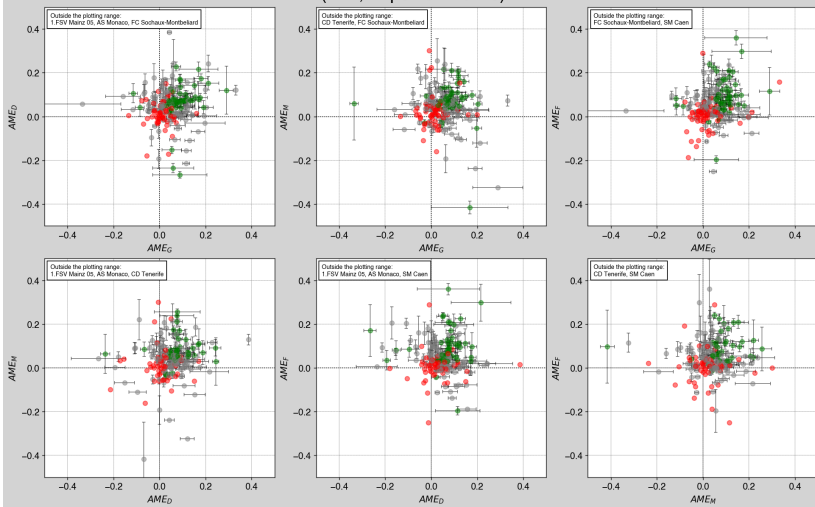
Average Marginal Effects across Positions (ALL, Expected Points)



► AMEs (Actual Points)

Heterogeneous Estimates: Pairwise AMEs I (Expected Points)

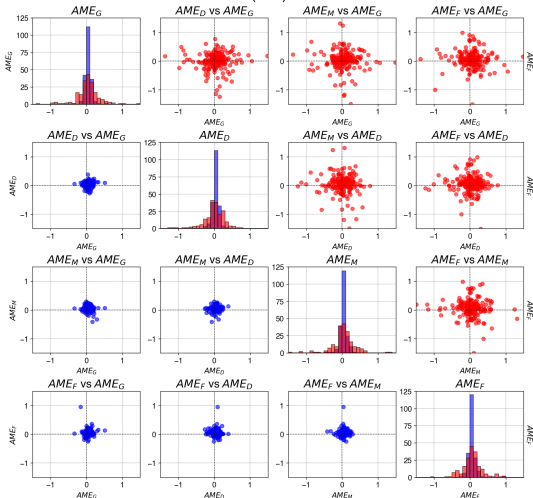
Pairwise Average Marginal Effects across Positions
(ALL, Expected Points)



► Pairwise AMEs I (Actual Points)

Heterogeneous Estimates: Pairwise AMEs II

Distribution of Average Marginal Effects
(ALL)

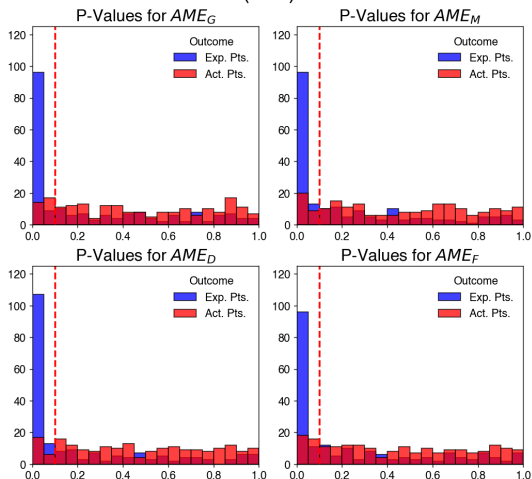


- Distribution of AMEs is similar across positions (around zero)
- No clear correlation across position groups
- Similar patterns hold for interaction-based γ 's

►► Pairwise γ 's

Heterogeneous Estimates: Distribution of p -Values

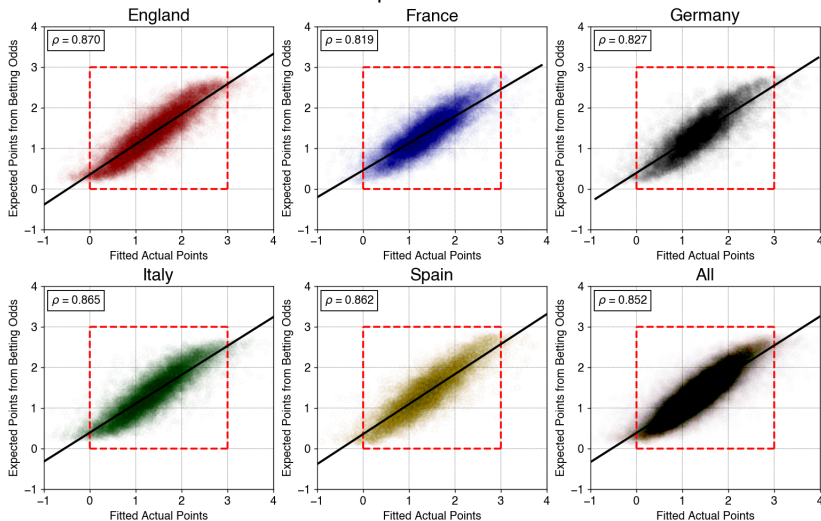
Distribution of P-Values for Average Marginal Effects (ALL)



- AMEs are positive and mostly significant for all positions with expected points as the outcome
- Uniform distribution for p -values; no significance beyond random chance with actual points

Heterogeneous Estimates: Fitted Points

Correlation between Expected Points and Fitted Points



Heterogeneous Estimates: Summary

- Average marginal effects represent increase in expected points after a log-unit increase in aggregate player valuation for the position group
- Dependent variable reflects **bookmakers' expectations** of team performance; results are no longer significant under actual points
 - Low significance with actual pts, **but** fitted values match expected pts
 - Reasonable due to intrinsic uncertainty in match outcomes; bookmakers' expectations approximate infinitely-repeated matches
 - Expected points “purge” noise from random in-match events
- Teams should recruit (transfer out) players in positions with highest (lowest) AMEs to improve performance on average
 - Heterogeneity across AMEs encourages transfer activity and endogenously creates variation in player valuations

Conclusion

Conclusion: Next Steps I

- **End Goal:** Construct a framework for player valuations and the Big 5 clubs' transfer activity in European football
 - Estimate mapping from player valuations to clubs' performances to obtain clubs' efficiency (AMEs with respect to positions)
 - Estimate mapping between player performances and valuations to obtain players' efficiency (AMEs with respect to statistics)
 - Build structural matching framework where clubs transfer players based on varying efficiencies and aging-related dynamics
- **Contribution:** Mobility with varying worker and firm productivity; possible extensions into academia (requires proxy for “valuation”)

Conclusion: Next Steps II

- Perform reduced form analysis to establish link between player performances and market valuations as well as wages
 - Re-estimate with rolling windows to allow time-varying AMEs,
 - Estimate quantile regressions for valuations on updated data,
 - Model finitely-lived asset valuation dynamics
- Write and estimate a structural model of the transfer market
 - ① Players are exogenously “found” by each team across Big 5,
 - ② Buyer approaches seller to negotiate transfer for a fee,
 - ③ Buyer negotiates contract with prospective signing,
 - ④ If agreed, new contract is written and previous contract amortized

Appendix

Revenues by League and Division

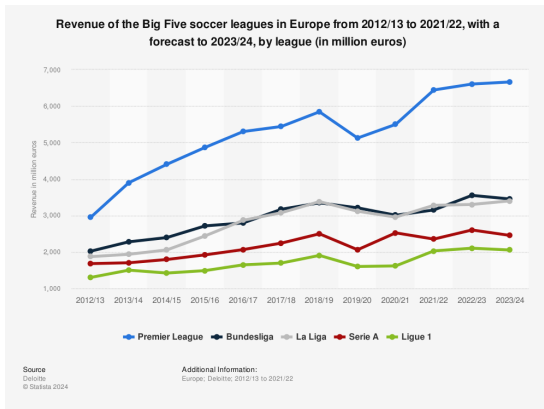


Figure 3: Revenues across Big 5 Major Leagues until 2022/23 (Statista)

► Competition NB: 2023 2nd Division revenues were €800m+ in England and Germany, €300m+ in Italy and Spain, and €150m+ in France. Eredivisie (NED) earns €500m+.

Revenues Compositions

Chart 2: 'Big five' European league clubs' revenue - 2021/22 (€m)

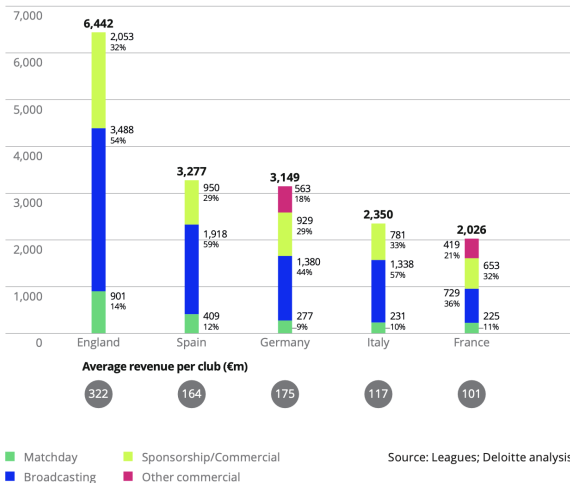
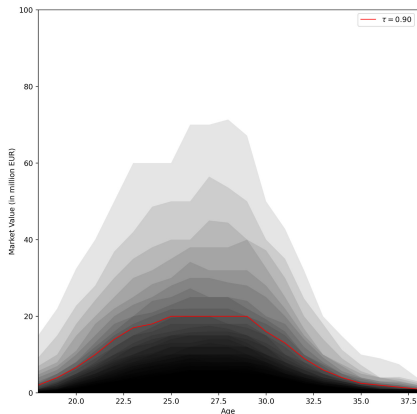


Figure 4: Revenues Composition across Big 5 in 2023 (Deloitte)

Transfer Market and Finance-based Regulations

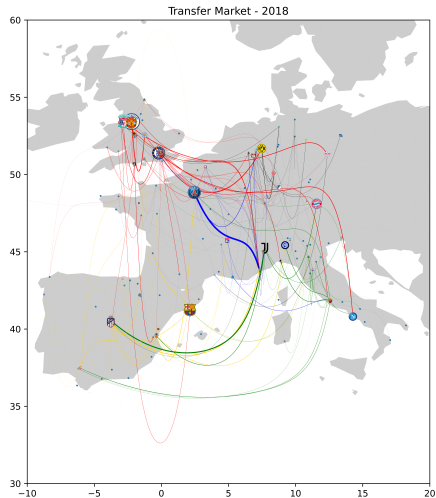
- Players are the main productive asset (and expense) for clubs
 - May be “grown” within the youth development system or acquired by paying a transfer fee to the current club and negotiating a new contract with the player
- Multiple foreign-based acquisitions of clubs located in London, Paris and other prominent European cities in early 2000's
 - Leagues (and major incumbents) claim “financial doping”; player transfers at significant premiums paid by owners
- UEFA introduces Financial Fair Play (FFP) to tie clubs' maximum allowed player-based expenses to club revenues
 - Only applicable in UEFA competitions, but similar rules exist (e.g. Profit and Sustainability Rules in England)

Distribution of Player Valuations



- Market value by age reflects peak in late-20s then decline
 - Aging dynamics contribute to regular transfer activity
- Ridge-plot shows the distribution of player valuations at each %tile for a given age
 - Highly skewed distribution; 90th %ile peaks at €20m, 99th %ile peaks at €70m

Transfer Market in 2018



Major transfers include:

- Kylian Mbappé (€135m)
ASM (FRA) → PSG (FRA)
- Cristiano Ronaldo (€100m)
RMA (ESP) → JUV (ITA)
- Naby Keita (€60m)
RBL (GER) → LIV (ENG)

Summary Statistics

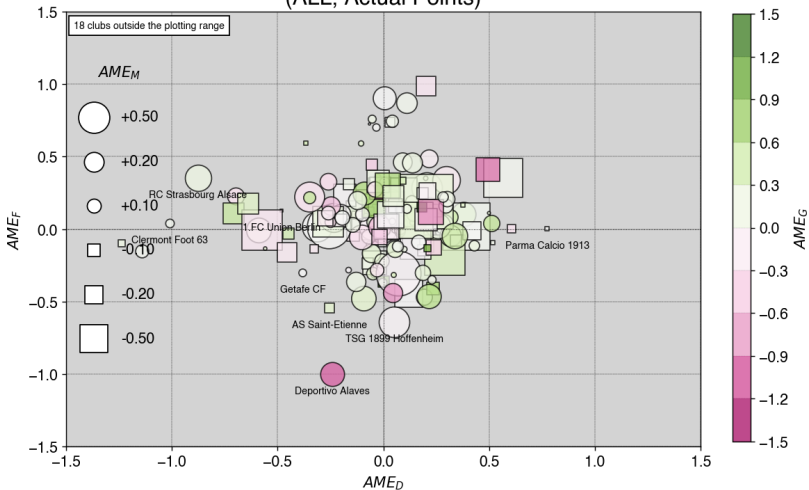
Table 2: Summary Statistics

	England				France				Germany				Italy				Spain				All			
	Mean	Std	Min	Max	Mean	Std	Min	Max	Mean	Std	Min	Max	Mean	Std	Min	Max	Mean	Std	Min	Max	Mean	Std	Min	Max
Expected Points	1.38	0.59	0.13	2.81	1.36	0.40	0.13	2.82	1.38	0.53	0.12	2.84	1.37	0.55	0.15	2.79	1.37	0.56	0.11	2.94	1.37	0.55	0.11	2.84
Actual Points	1.38	1.32	0	3	1.36	1.30	0	3	1.38	1.32	0	3	1.37	1.31	0	3	1.38	1.32	0	3	1.37	1.31	0	3
Goalkeepers	€10.89m	€13.33m	€0.05m	€90.00m	€4.89m	€5.75m	€0.05m	€65.00m	€5.48m	€6.28m	€0.05m	€45.00m	€6.36m	€7.93m	€0.03m	€60.00m	€8.57m	€14.43m	€0.05m	€100.00m	€7.32m	€10.58m	€0.03m	€100.00m
Mean Defenders	€11.15m	€10.63m	€0.42m	€81.25m	€4.28m	€5.16m	€0.19m	€56.25m	€7.27m	€8.18m	€0.43m	€71.67m	€6.40m	€7.57m	€0.08m	€56.67m	€7.29m	€8.81m	€0.21m	€62.50m	€7.30m	€8.57m	€0.08m	€81.25m
Mean Midfielders	€15.20m	€14.72m	€0.30m	€110.00m	€5.85m	€6.86m	€0.25m	€70.00m	€8.25m	€10.15m	€0.33m	€85.00m	€8.71m	€8.81m	€0.20m	€65.00m	€9.69m	€12.71m	€0.40m	€97.50m	€9.62m	€11.50m	€0.20m	€110.00m
Mean Forwards	€19.51m	€21.00m	€0.10m	€180.00m	€7.61m	€12.98m	€0.15m	€180.00m	€10.31m	€12.86m	€0.08m	€150.00m	€11.61m	€12.67m	€0.05m	€105.00m	€13.10m	€18.68m	€0.05m	€136.67m	€12.56m	€16.66m	€0.05m	€180.00m
Min Defenders	€5.14m	€6.22m	€0.05m	€56.00m	€1.78m	€2.56m	€0.05m	€45.00m	€3.18m	€4.69m	€0.03m	€65.00m	€2.67m	€4.15m	€0.03m	€55.00m	€3.02m	€4.82m	€0.03m	€50.00m	€3.16m	€4.78m	€0.03m	€65.00m
Min Midfielders	€7.64m	€9.04m	€0.05m	€80.00m	€2.92m	€4.25m	€0.04m	€65.00m	€4.13m	€6.86m	€0.03m	€70.00m	€3.86m	€5.38m	€0.05m	€60.00m	€4.65m	€7.39m	€0.05m	€85.00m	€4.67m	€6.90m	€0.03m	€85.00m
Min Forwards	€12.51m	€15.84m	€0.05m	€180.00m	€4.17m	€6.68m	€0.05m	€180.00m	€6.15m	€9.10m	€0.03m	€150.00m	€6.93m	€9.07m	€0.05m	€100.00m	€7.15m	€11.92m	€0.03m	€180.00m	€7.45m	€11.42m	€0.03m	€180.00m
Max Defenders	€17.89m	€16.45m	€0.50m	€110.00m	€7.39m	€9.22m	€0.25m	€75.00m	€12.18m	€12.65m	€1.00m	€85.00m	€11.30m	€13.52m	€0.10m	€75.00m	€12.76m	€14.83m	€0.40m	€90.00m	€12.32m	€14.02m	€0.10m	€110.00m
Max Midfielders	€24.05m	€23.43m	€0.30m	€150.00m	€9.39m	€11.18m	€0.40m	€180.00m	€13.44m	€16.25m	€0.59m	€120.00m	€15.19m	€15.50m	€0.45m	€110.00m	€15.32m	€19.89m	€0.50m	€180.00m	€15.60m	€18.44m	€0.30m	€180.00m
Max Forwards	€27.11m	€29.38m	€0.10m	€180.00m	€11.73m	€22.74m	€0.15m	€200.00m	€14.86m	€18.94m	€0.09m	€150.00m	€16.75m	€18.61m	€0.05m	€120.00m	€20.29m	€28.91m	€0.05m	€180.00m	€18.32m	€24.87m	€0.05m	€200.00m
4-3-3 Formation	42.71%	42.71%	0%	100%	49.56%	49.56%	0%	100%	35.41%	35.41%	0%	100%	30.52%	30.52%	0%	100%	51.82%	51.82%	0%	100%	42.17%	42.17%	0%	100%
4-4-2 Formation	27.30%	44.55%	0%	100%	19.58%	39.68%	0%	100%	26.04%	43.88%	0%	100%	25.22%	43.43%	0%	100%	15.30%	36.00%	0%	100%	22.60%	41.82%	0%	100%
4-2-4 Formation	7.63%	26.55%	0%	100%	9.26%	28.98%	0%	100%	14.81%	35.52%	0%	100%	2.73%	16.29%	0%	100%	14.43%	35.14%	0%	100%	9.55%	29.40%	0%	100%
3-5-2 Formation	4.75%	21.28%	0%	100%	4.73%	21.23%	0%	100%	9.01%	28.64%	0%	100%	25.82%	43.77%	0%	100%	2.65%	16.05%	0%	100%	9.48%	29.29%	0%	100%
4-5-1 Formation	6.82%	25.21%	0%	100%	7.82%	26.85%	0%	100%	5.29%	22.38%	0%	100%	5.89%	23.48%	0%	100%	4.21%	20.99%	0%	100%	6.02%	23.79%	0%	100%
5-3-2 Formation	2.51%	15.63%	0%	100%	3.33%	17.94%	0%	100%	1.33%	11.49%	0%	100%	2.16%	14.52%	0%	100%	4.22%	20.11%	0%	100%	2.75%	16.36%	0%	100%
3-6-1 Formation	2.93%	18.85%	0%	100%	1.83%	13.40%	0%	100%	3.61%	18.65%	0%	100%	3.65%	18.76%	0%	100%	1.11%	10.46%	0%	100%	2.60%	15.90%	0%	100%
3-4-3 Formation	1.75%	13.13%	0%	100%	1.71%	12.98%	0%	100%	3.15%	17.48%	0%	100%	3.44%	18.22%	0%	100%	1.45%	11.95%	0%	100%	2.28%	14.91%	0%	100%
5-2-3 Formation	2.57%	15.83%	0%	100%	1.55%	12.37%	0%	100%	1.08%	10.34%	0%	100%	0.29%	5.34%	0%	100%	2.71%	16.22%	0%	100%	1.66%	12.77%	0%	100%
Other Formation	1.02%	10.07%	0%	100%	0.62%	7.88%	0%	100%	0.27%	5.21%	0%	100%	0.32%	5.61%	0%	100%	2.11%	14.39%	0%	100%	0.89%	9.40%	0%	100%

► Homogeneous Estimates

Heterogeneous Estimates: AMEs (Actual Points)

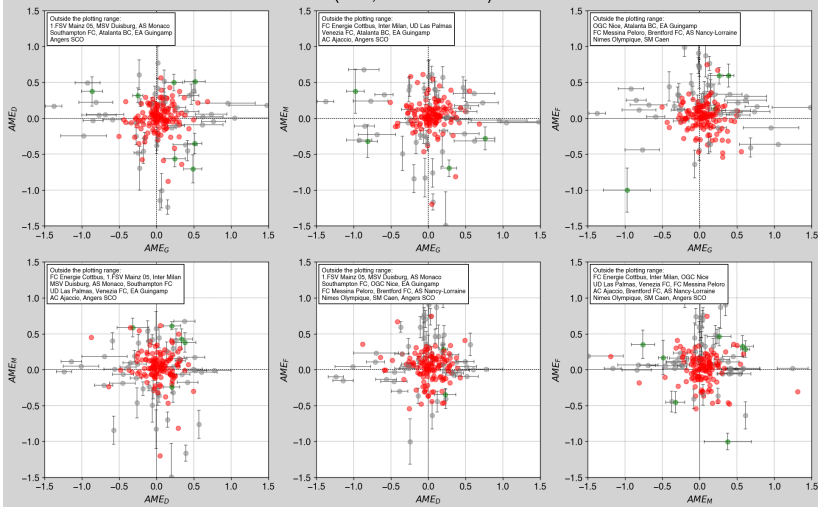
Average Marginal Effects across Positions
(ALL, Actual Points)



▶ AMEs (Expected Points)

Heterogeneous Estimates: Pairwise AMEs I (Actual Points)

Pairwise Average Marginal Effects across Positions
(ALL, Actual Points)



► Pairwise AMEs I (Expected Points)

Heterogeneous Estimates: Pairwise γ 's

Distribution of Interaction Terms γ ,
(ALL)

