



What Wins the EPL Trophy?



Player Quality, Tactics, or Money? Decomposing English Premier League Success

Ayush Pundir | 24507, BTech CSE(DS), IIIT Una, Himachal Pradesh, India

Dr. Prince Sharma | Faculty, SoC, Indian Institute of Information Technology Una, Himachal Pradesh, India

KEY FINDINGS: ✓ **Player quality is dominant ($R^2 = 0.879$).** : ✗ **Tactical style adds nothing once quality is controlled.** ✗ **Transfer spend has no independent effect ($R^2 = 0.012$).**

1 Introduction



Does EPL success come from player quality, tactical style, or transfer spend? We test three factors across 100 team-seasons, five EPL seasons (2019/20–2023/24), using block OLS regression + incremental F-tests. All analysis is built in R; visualisations use ggplot2.

Factors Tested

- Player Quality (npxGD, Finishing)
- Tactical Style (PPDA, Deep Completions)
- Transfer Spend (Gross, Net)

Why 2019/20–2023/24?

- 2019/20: Liverpool's title; COVID stoppage & restart
- 2020/21: Man City champions; matches behind closed doors
- 2021/22: Man City's +73 GD; tightest title race in the sample
- 2022/23: World Cup mid-season break; squad depth tested
- 2023/24: Man City's 4th straight title; Arsenal 2 pts behind

2 Data & Methods



R packages: tidyverse · broom · ggplot2 · patchwork · ggrepel

- Data Sources**
Match results (football-data.co.uk) + xG/tactics (Understat) + transfer spend (Transfermarkt), all via Kaggle
- Reproducibility Fix**
FBref scraping blocked (HTTP 403) — pivoted to Kaggle-hosted datasets; fully scripted and reproducible
- Variable Construction**
npxGD (not raw GD — avoids tautology with points), finishing overperformance, PPDA, deep completions, gross/net spend — standardised
- Block OLS Models**
4 models via lm(): 3 single-block (Quality / Tactics / Money) + 1 full model with all blocks
- F-Tests & Visuals**
Incremental F-tests via anova(); plots built with ggplot2 + patchwork

3 Conclusions



- ✓ **Player Quality Dominates**
npxGD + finishing overperformance significant at $p < 0.001$ in every specification
- ✗ **Tactics Not Independent**
Adds nothing once quality is controlled ($F = 1.49$, $p = 0.231$)
- ✗ **Money Not Significant**
No independent effect on points ($F = 0.10$, $p = 0.903$)
- ★ **Full Model**
 $R^2 = 0.883$ — quality alone explains nearly all of it; $n = 100$ team-seasons

4 References & Data



- [1] Understat game-level data — via Kaggle (codytipton/understat-data)
- [2] Transfermarkt player-scores — via Kaggle (davidcariboo/player-scores)
- [3] football-data.co.uk match archives
- [4] Szymanski (2010) · Caley (2015) · Mackay (2017)

5 Acknowledgments



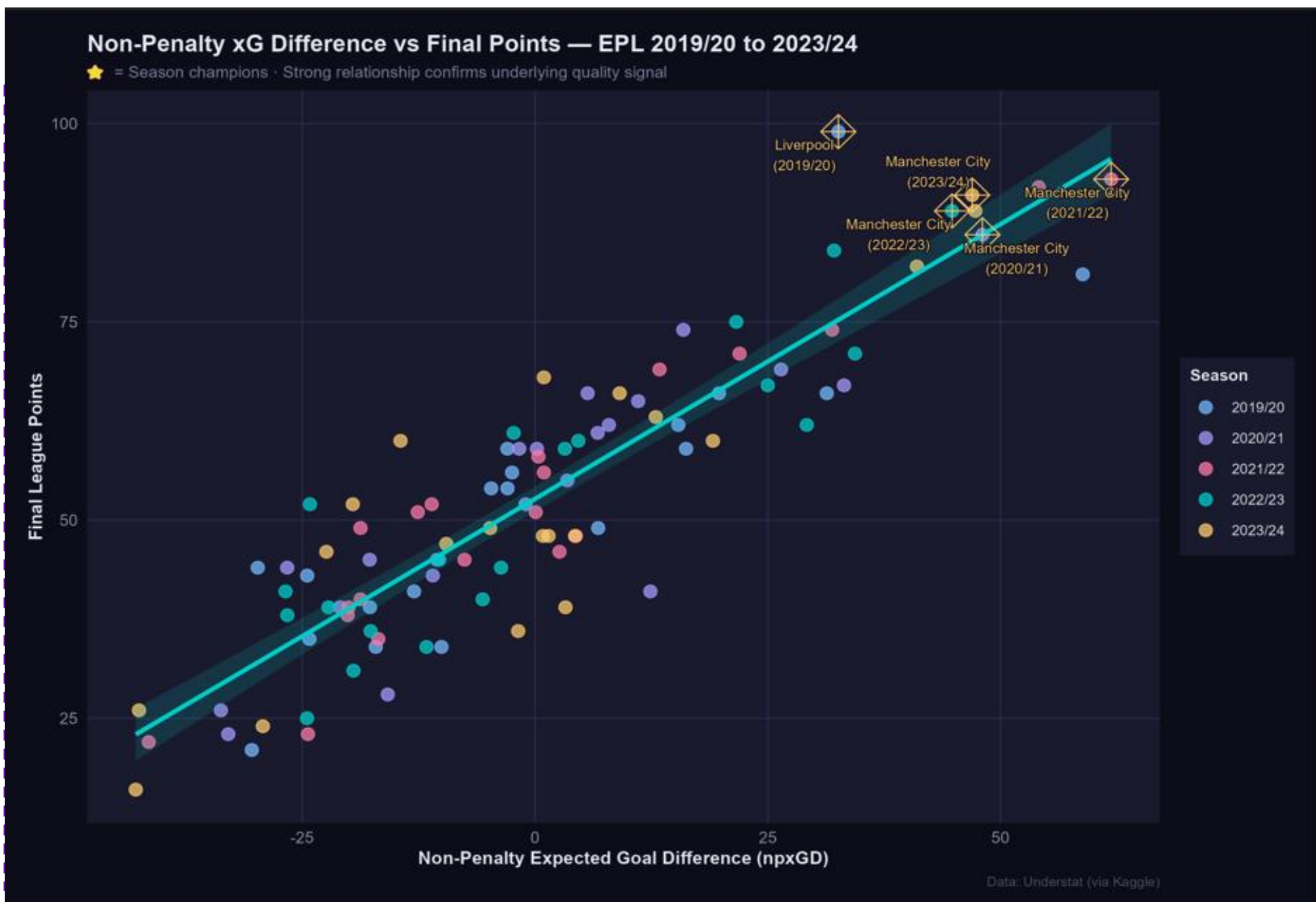
Thanks to the R and useR! community for tools and open-source support, and to Kaggle dataset maintainers for public data access.

01 Player Quality — DOMINANT ✓

0.879

R^2 · single-block OLS

Non-penalty xG difference (npxGD) and finishing overperformance. Both significant at $p < 0.001$ in the full model ($\beta = 13.7$ and $\beta = 4.8$).



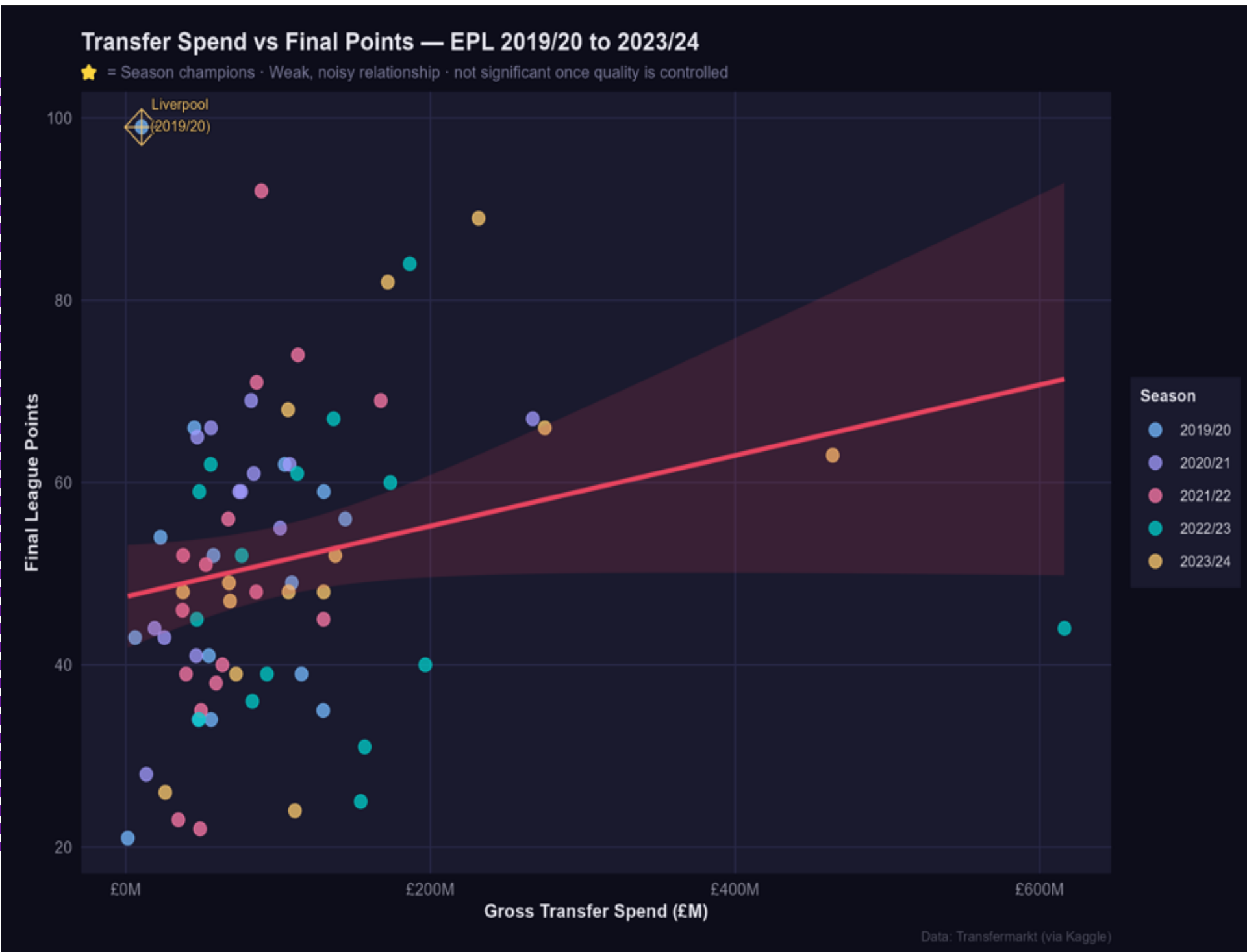
The strongest and only unambiguous predictor of EPL points across all specifications.

03 Transfer Spend — NOT SIGNIFICANT ✗

0.012

R^2 · single-block OLS (adj. $R^2 < 0$)

Gross and net transfer spend. Neither approaches significance in the full model ($p = 0.76$, $p = 0.93$).

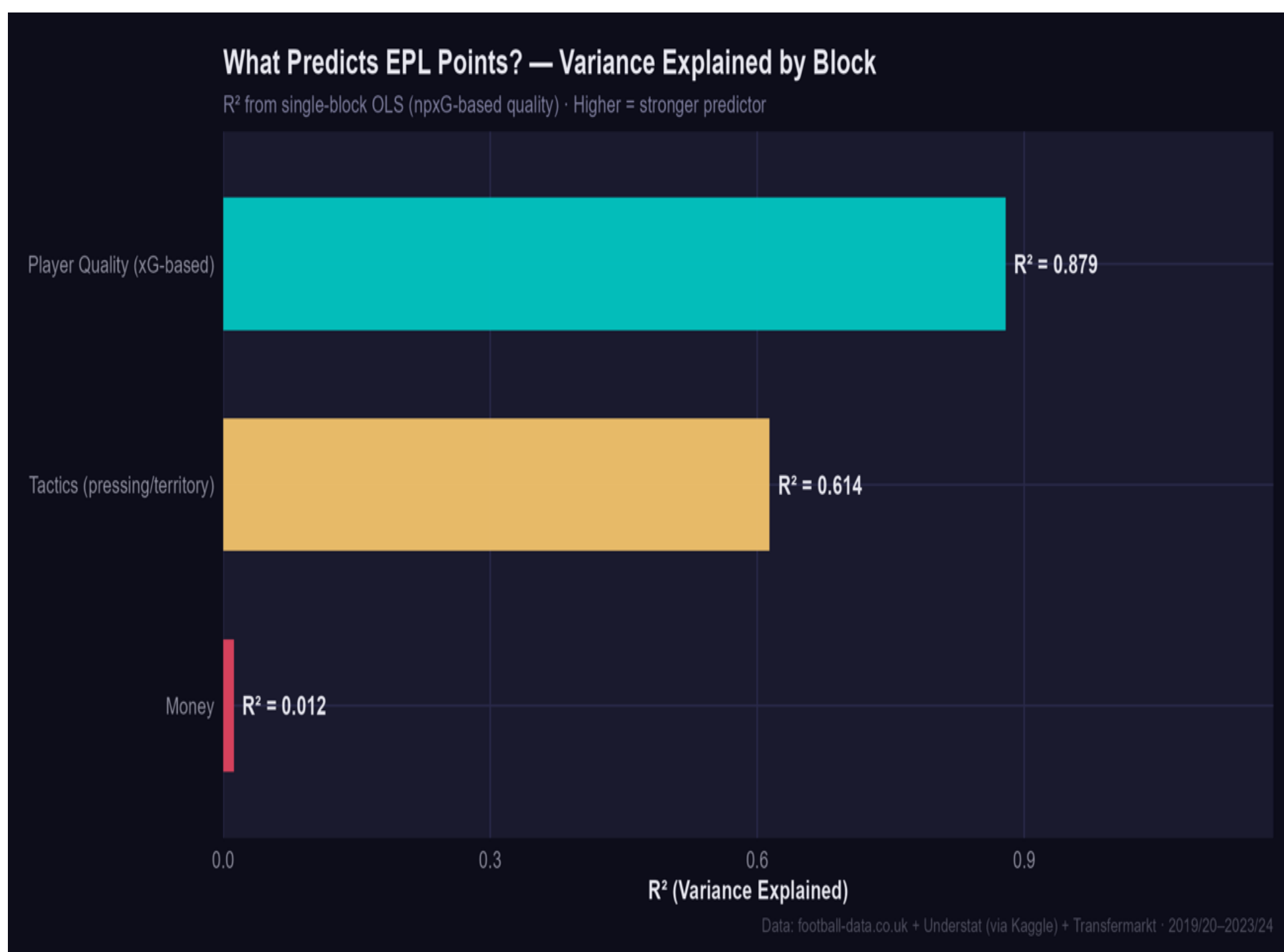


F-test: adding money to quality + tactics does NOT improve fit ($F = 0.10$, $p = 0.903$).

05 R^2 by Block — DOMINANT ✓

Quality vs Tactics vs Money vs Full

Quality: 0.879 (Adj. 0.876) · Tactics: 0.614 (Adj. 0.606) · Money: 0.012 (Adj. -0.008) · Full: 0.883 (Adj. 0.875)



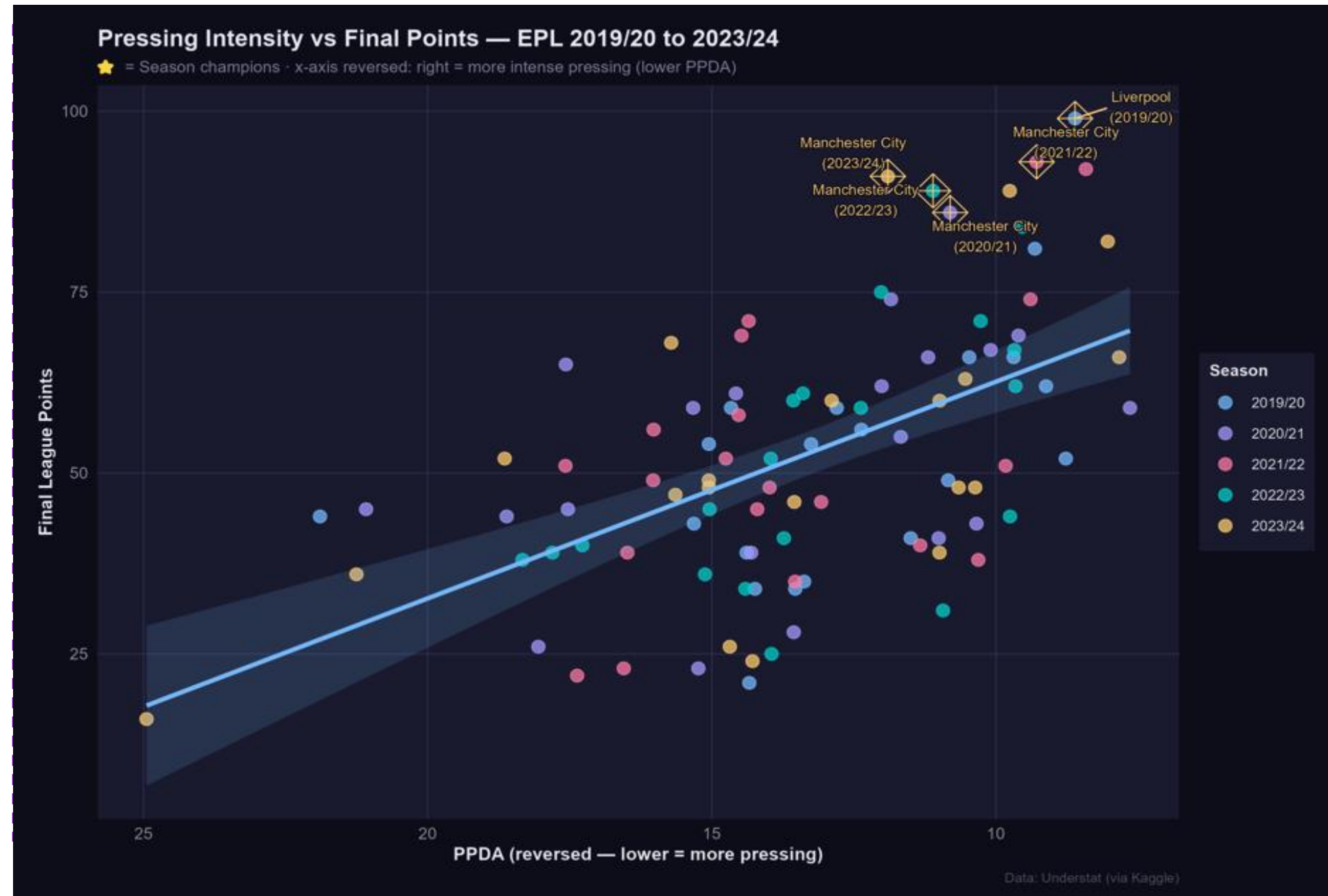
Quality alone nearly matches the full model — tactics and money add almost nothing on top.

02 Tactical Style — NOT INDEPENDENT ✗

0.614

R^2 · single-block OLS

Pressing intensity (PPDA) and territorial control (deep completions). Neither significant in the full model ($p = 0.75$, $p = 0.08$).



F-test: adding tactics to quality alone does NOT significantly improve fit ($F = 1.49$, $p = 0.231$).

04 Full Model — CONSISTENT ✓

0.883

R^2 · all blocks combined (Adj. 0.875)
Standardised coefficients with 95% CI, OLS (lm()). Only quality-block predictors (npxGD, finishing overperformance) are significant.



$n = 100$ team-seasons · $F(6,93) = 116.5$, $p < 0.001$

06 Champion Profile — CONSISTENT ✓

Champions vs rest of league, 5-season average
Champions lead on every metric, but the gap is largest for npxGD and finishing — not transfer spend.



Three of five champions recorded minimal net transfer spend in their title-winning season.