Journal of High School Science



What possession is worth: evidence from two world cups

Sultan R

Submitted: July 25, 2026, Revised: version 1, August 29, 2026, version 2, October 12, 2026, version 3, October 18, 2026
Accepted: October 24, 2026

Abstract

At the 2022 FIFA World Cup, the team with more of the ball won 23 of 49 decided matches; at the 2026 World Cup, it won 61 of 80. This study asks why. FIFA post-match statistics for all 168 matches were combined with pre-tournament Elo ratings and modeled with one observation per match. Possession tracked team strength almost equally in both tournaments ($r = 0.61$ and 0.62) and, once strength was known, added no information about goal difference (2022: $b = -0.024$, $p = 0.42$; 2026: $b = +0.004$, $p = 0.79$). On the win-draw-loss scale, strength-adjusted possession was negatively associated with results in 2022 (ordinal $b = -0.087$, $p = 0.002$) but not in 2026, consistent with the 2022 upsets, in which successful underdogs had substantially less possession and more transition play than unsuccessful underdogs. Ablation and double machine learning located possession's information: in 2026 matches with a clear favorite, adding possession to strength and directness increased explained variance by 0.54 for completed line breaks, 0.14 for expected goals (xG), and only 0.013 for goal difference. Cross-fitted estimates indicated approximately 15 additional line breaks and +0.15 xG, but no statistically distinguishable change in goal difference, per ten percentage points of possession. Defensive denial accounted for approximately 43% of the positive xG volume component, while lower goal production per unit of possession offset the associated scoring opportunities. These findings distinguish possession's role as an enabling resource from its productivity and incremental predictive significance: a resource can be fundamental to attacking production without explaining additional variation in realized goals. This distinction motivates *Sublime Football* - a proposed extension of *Total Football* - consisting of possession-oriented tactical philosophy in which possession is evaluated by its conversion into attacking production, defensive denial, and scoring advantage rather than by its quantity alone. The framework remains a hypothesis for future empirical testing.

Keywords

Football analytics, FIFA World Cup, Ball possession, Team strength, Expected Goals (xG), Attacking efficiency, Tactical analysis, Counterattacking, Possession productivity, Total Football

Robert Sultan, Rochambeau The French International School, 9600 Forest Road, Bethesda, MD 20814, USA and Statistics, Montgomery College, 51 Mannakee Street, Rockville, MD 20850, USA. robertrsultan@gmail.com

https://doi.org/10.64336/001c.173299

1. Introduction

Two consecutive World Cups yielded opposite answers to the same question. At Qatar 2022 the team with more possession of the ball won 23 of 49 decided matches (46.9%); at the first 48-team World Cup in 2026 it won 61 of 80 (76.2%). The difference between those proportions is too large to attribute comfortably to chance ($z = 3.39$, $p < 0.001$). Read side by side, the two tournaments suggest that the value of possession depends on the tournament, which describes the result without explaining it.

Three explanations make different predictions. The first is team quality: stronger teams both keep the ball and win, so possession may mark the stronger side rather than help it. The second is strategy: the same share of the ball can represent productive positional attack, harmless circulation, or a team chasing the score, and low possession can be deliberate counter-attacking or simple inability to keep the ball. The third is chance creation: possession may matter only to the extent that it produces shooting opportunities. These explanations are not mutually exclusive, and an observational study cannot separate them causally, but they can be distinguished statistically when quality, style and chance creation are measured independently.

The literature is more sceptical of possession than popular commentary. Collet (1) found that ball retention predicts league position but is a weak predictor of individual matches once team quality is considered, and FIFA's technical material and recent analysis of World Cup scoring emphasise chance creation over time on the ball (2,3). What the two most recent World Cups add is unusually complete measurement: for every match FIFA published a post-match report giving possession, passing, line breaks, receptions in the final third, pressing, forced turnovers and the share of each team's time spent in defined phases of play, with expected goals (xG) added in 2026 (4,5).

This paper uses those reports to ask why possession appeared to change its value between 2022 and 2026. It does not treat a vanishing coefficient as proof that possession carries no information. Instead it separates possession's total association with results from the part shared with team strength, progression and chance creation; tests directly whether the tournaments differed; and examines the matches in which the stronger team failed to win, since those matches turn out to carry the cross-tournament difference. Finally, it asks where possession's information lies, separating territory, chance quality and goals, with an ablation analysis and double machine learning.

2. Materials and Methods

2.1 Data sources and sample

The sample is every match of the 2022 and 2026 FIFA World Cups: 64 and 104 matches, 168 in total. Team statistics were extracted from FIFA's official post-match summary report for each match (4,5), a PDF published on the FIFA Training Centre, using the Poppler pdftotext utility (version 26.07) and regular expressions applied to the key-statistics and phases-of-play pages. All 168 reports were retrieved and parsed; no match was excluded.

Expected goals appear only in the 2026 reports. Results, extra time, penalty shoot-outs, goal minutes and penalty goals were taken from the openfootball World Cup dataset (6) and checked against the final scores printed in each FIFA report; all 168 agreed. Red cards were taken from the line-up tables of the Wikipedia match pages (7); a dismissal after the final whistle (Netherlands v Argentina, 2022) is not counted. Elo ratings were extracted from the World Football Elo Ratings archive (8).

An earlier version of this analysis reported 62 matches for 2022. Two reports, Spain versus Costa Rica and Germany versus Japan, are stored by FIFA under a different web directory and were silently missed by the earlier download. Both are included here. Table 2 reconciles every count used in the paper.

Team strength is each team's Elo rating immediately before its first match of the tournament, which is the rating on the eve of the tournament because no team plays a match between the opening day and its own first fixture. These ratings precede every tournament result and so cannot absorb the outcomes being explained. Two alternatives are used as sensitivity analyses: the end-of-year rating from the previous December, as in the earlier version, and each team's rating immediately before each individual match. The predictor is the Elo difference between the two teams, per 100 points. Tournament hosts (Qatar in 2022; Canada, Mexico and the United States in 2026) are identified by a host indicator.

2.2 Variables

Table 1 defines every variable. No team was assigned a style label; every tactical variable is either a FIFA phase-of-play percentage, computed by FIFA from tracking data, or a fixed formula applied to counts in the FIFA report. Variable definitions were fixed before outcomes were modelled. That is not a substitute for prospective preregistration, and analytic flexibility is addressed instead by reporting every specification tested, with multiplicity adjustment.

Table 1. Variables, definitions and sources. Every model uses the difference between the two teams.

Variable	Definition	Source
Goal difference (GD)	Team A goals minus team B goals, including extra time; shoot-outs count as draws	FIFA report; openfootball (6)
Result	Win, draw or loss for team A on the same basis	As above
Possession	Team A's share of the two-team possession total minus 50 (percentage points)	FIFA report
Elo gap	Team A minus team B Elo rating before the tournament, per 100 points	eloratings.net (8)
Host	Host indicator, team A minus team B	Tournament records
Pass completion	Completed passes as % of passes	FIFA report
Line breaks rate	Completed line breaks per 100 passes	Computed from FIFA report
Final-third rate	Receptions in the final third per 100 passes	Computed from FIFA report

Counter-attack, attacking transition, long ball	% of in-possession time in each phase	FIFA report
High press, low block	% of out-of-possession time in each phase	FIFA report
PPDA	Opponent passes per defensive pressure applied	Computed from FIFA report
Forced turnovers	Count	FIFA report
Shots, shots on target	Attempts at goal and attempts on target	FIFA report
Directness; shots on target rate	Shots, or shots on target, per 100 passes	Computed from FIFA report
Expected goals (xG)	Sum of chance quality; 2026 only	FIFA report
Scored first	+1 if team A scored the opening goal, −1 if team B did, 0 if goalless	openfootball (6)
Red cards; penalty goals	Counts	Wikipedia (7); openfootball (6)

2.3 Statistical approach

The unit of analysis is the match. For each match, team A is the team listed first by FIFA, every predictor is the difference between team A and team B, and the outcome is team A's goal difference. Swapping the labels reverses the sign of every variable and leaves the fitted model unchanged, so models are fitted without an intercept; for the same reason a tournament main effect is zero by construction, and tournament enters only through interactions. The basic model is given by,

$$GD_i = \beta_1 \Delta Elo_i + \beta_2 Poss_i + \beta_3 \Delta Host_i + \gamma' \Delta Z_i + \varepsilon_i \quad \text{(eq1)}$$

In equation (eq1), Poss is team A's possession minus 50 percentage points (half the possession gap) and ΔZ are further differenced covariates. Standard errors are heteroscedasticity-robust (HC3). The earlier version used two mirrored observations per match with standard errors clustered by match; Section 3.1 shows that for possession and Elo the two formulations give identical point estimates, as they must algebraically, but that they diverge for tactical variables, because a team's own counter-attack share is not the negative of its opponent's. All main analyses therefore use one row per match.

The analyses below were specified in full before the models reported here were fitted, although none was preregistered. (i) Nested models fitted separately to each tournament and to the pooled data: M0, Elo gap and host; M1, adding possession; M2, adding ten tactical variables (pass completion, line-break rate, final-third rate, counter-attack, attacking-transition, long-ball, high-press and low-block shares, PPDA and forced turnovers); M3, adding shots and shots on target; and in 2026 M4, adding xG. Each step reports the possession coefficient with its 95% confidence interval, adjusted R^2, a nested F test, and repeated ten-fold cross-validated root mean squared error (RMSE; 50 repeats). Because shots and xG may lie downstream of possession, M1 estimates possession's total association given quality, and M3 and M4 estimate only the residual association once chance creation is held fixed. (ii) A pooled model with Elo × tournament and possession × tournament interactions. (iii) Ordinal logistic (loss, draw, win) and logistic (decided matches only) models, and 90-minute goal difference, as outcome-scale sensitivity analyses. (iv)

Residualised possession: possession predicted out of fold (five-fold cross-fitting, 50 repeats) from Elo, host, directness, pass completion, phase-of-play shares, PPDA and which team scored first, with and without passing volume; the out-of-fold residual is the possession a team held beyond what its quality and style predicted. (v) Possession × strategy interactions, formed from team-level centred products so that they remain valid differences, with Benjamini-Hochberg false-discovery-rate and Holm adjustment and an omnibus F test. (vi) Possession efficiency (per-100-pass rates) and, equivalently, possession × chance-production interactions. (vii) Four team-performance profiles formed by possession above or below 50% and shots above or below the tournament median. (viii) A comparison of the frequency of each style across tournaments. (ix) An upset analysis. (x) Commonality and general dominance analysis of R^2. (xi) Principal components of team style with varimax rotation. (xii) Elastic-net regression with bootstrap retention frequencies. (xiii) A path decomposition with match-level bootstrap confidence intervals. (xiv) Elo-matched and inverse-probability-weighted comparisons. Two further analyses ask where possession carries information, separately for territory, chance quality and goals. They use the 2026 matches with a clear favourite, an Elo gap of at least 50 points as in the upset analysis, and orient every variable as underdog minus favourite, with an intercept in every model. Possession enters as the possession difference in percentage points, scaled per ten points, and directness is shots per 100 completed passes. Five outcomes are modeled as underdog-minus-favourite differences in completed line breaks, receptions in the final third, xG, goals and forced turnovers. (xv) Ablation: for each outcome a baseline model containing the Elo gap and directness is compared with the same model plus possession, on the same matches, by the gain in R^2, a nested F test and repeated ten-fold cross-validated RMSE (50 repeats). A targeted ablation for xG then adds line breaks and final-third receptions before possession, to test whether territorial progression exhausts possession's information about chance quality. (xvi) Double machine learning (DML) (10) estimates the partially linear model given by,

$$Y_i = \theta T_i + g(X_i) + \varepsilon_i, \quad T_i = m(X_i) + \nu_i \quad \text{(eq2)}$$

in which T is the possession difference, Y one of four outcomes (line breaks, final-third receptions, xG and goal difference) and X the adjustment set. The nuisance functions g and m are fitted out of fold by five-fold cross-fitting, and θ is estimated by regressing the out-of-fold residual of Y on the out-of-fold residual of T. Two learners are used: ridge regression, a transparent benchmark whose penalty is chosen by cross-validation inside each training fold, and a random forest (200 trees, minimum leaf size 5) to allow nonlinear adjustment. Cross-fitting is repeated 50 times with different fold splits; the reported estimate is the median over repetitions and its standard error combines the within-split variance with the spread across splits (10). The primary adjustment set contains only pre-match variables: the Elo gap and

whether the underdog was FIFA's designated home team. Two further sets add line breaks and final-third receptions, and then also directness. Because these are measured during the match and may lie on the path from possession to chances and goals, the conditional specifications are sensitivity analyses, not estimates of possession's total effect. DML estimates are compared with ordinary least squares on the same covariates. Under the assumptions of no unmeasured confounding, correct temporal ordering and overlap, θ would be a causal effect; none of these can be defended with whole-match aggregates, so θ is interpreted throughout as an adjusted association. As robustness checks, both analyses were repeated with the host indicator in place of home designation, and for all 104 matches of 2026, all 64 of 2022 and the pooled data in the team-A-minus-team-B design of equation (eq1), with the host indicator in the baseline; in that design each training fold was augmented with its sign-reversed mirror image, so that the fitted functions reverse sign when the team labels are swapped.

Each match is classified as Elo-confirming, a draw, or an upset, from the favourite's point of view, with penalty shoot-outs counted as draws. Matches with an Elo gap below 50 points have no clear favourite and are excluded from the upset analysis; thresholds of 0 and 100 points are reported as sensitivity analyses. Upset and confirming matches are compared on underdog-minus-favourite differences using Mann-Whitney tests with false-discovery-rate adjustment. Because upsets and confirming results can occur at different Elo gaps, the contrasts are repeated in linear models adjusted for the gap, and a pooled logistic model of upset versus confirming result, with the Elo gap, tournament and their interaction, tests whether each mechanism differed by tournament. Forced-turnover counts differ so much between the two tournaments (means of 71 and 42 per team) that FIFA's definition appears to have changed, hence this variable is standardised within each tournament before any pooled analysis.

Two tournaments do not constitute a random sample of football, hence p-values are read as measures of how surprising a pattern would be under chance rather than as inferences to a wider population. Computation used Python 3.9 with pandas, NumPy, SciPy, statsmodels (9), scikit-learn and Matplotlib.

3. Results and Discussion

3.1 Sample and the descriptive contrast

Figure 1 shows the raw contrast. In 2022 the possession gap was weakly related to goal difference ($r = +0.15$, $p = 0.23$); in 2026 the relationship was clearer ($r = +0.41$, $p < 0.001$). The unadjusted possession coefficient was +0.023 (95% CI −0.028 to +0.074; $p = 0.37$) goals per percentage point in 2022 and +0.073 (95% CI +0.042 to +0.104; $p < 0.001$) in 2026. Comparing a significant coefficient with a nonsignificant one does not establish that they differ, and the direct test is less conclusive than the headline proportions: in a pooled model the unadjusted possession × tournament interaction is +0.050 (95% CI −0.009 to +0.110; $p = 0.10$).

The one-row and two-row formulations agree

where they should. With end-of-year Elo, the 2026 possession coefficient is +0.012 in both formulations, and the Elo coefficient is +0.458 in both. Similarly, in 2022, the possession coefficient is −0.035 in both formulations, and the Elo coefficient is +0.481 in both. For a non-complementary tactical variable they do not: the counter-attack coefficient in 2026 is +0.596 in the two-row model and +0.499 in the differenced model.

Table 2. Sample reconciliation and descriptive contrast.

Variable	2022	2026	Pooled
Matches (one row each)	64	104	168
Knockout matches	16	32	48
Went to extra time	5	9	14
Decided by penalty shoot-out (counted as draws)	5	4	9
Drawn (after 90 minutes or extra time)	15	24	39
Decided matches	49	80	129
Goalless matches	7	8	15
Higher-possession team won (decided)	23/49	61/80	84/129
Elo favourite won (decided)	35/49	69/80	104/129
r, possession gap and goal difference	+0.15	+0.41	+0.32
r, possession gap and Elo gap	+0.61	+0.62	+0.62
r, Elo gap and goal difference	+0.41	+0.62	+0.55

Decided means a winner after 90 minutes or extra time. All 2022 reports are included; the earlier version omitted two.

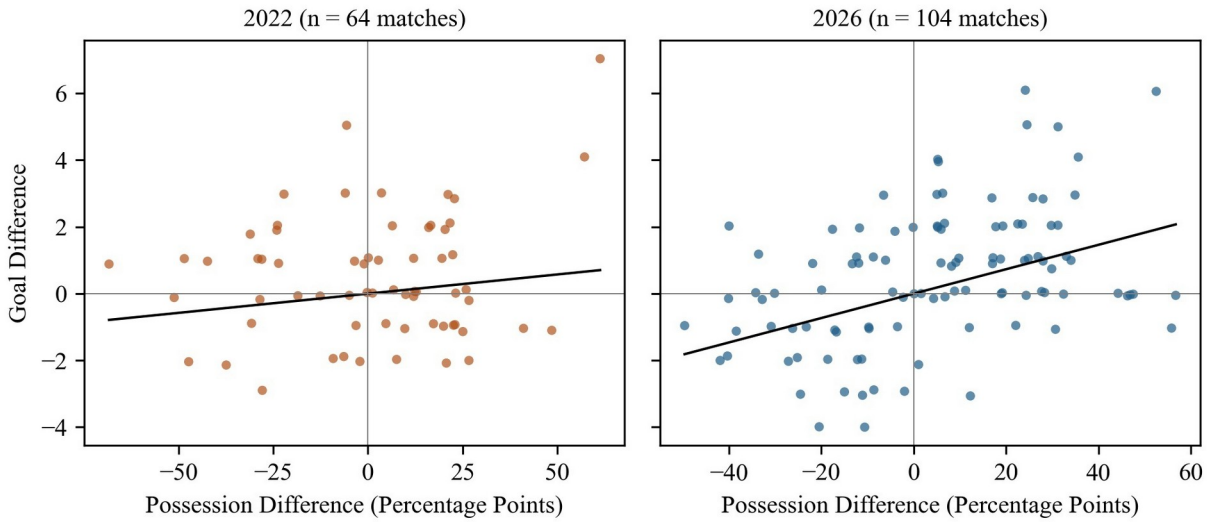


Figure 1. Possession difference and goal difference by tournament. Each point is one match (team A minus team B; small vertical jitter added). Lines are least-squares fits through the origin. 2022: 64 matches; 2026: 104 matches.

3.2 Possession adds no information about goal difference beyond team strength

Adding possession to a model containing the Elo gap and the host indicator (Table 3, M0 to M1) produced no improvement in either tournament. In 2022 the possession coefficient was −0.024 (95% CI −0.083 to +0.034; p = 0.42), adjusted R^2 changed by +0.006 and cross-validated prediction error increased by 0.042 goals; in 2026 it was +0.004 (95% CI −0.027 to +0.036; p = 0.79), adjusted R^2 changed by −0.005 and prediction error increased by 0.016. Possession correlated with the Elo gap almost identically in the two tournaments (r = +0.61 and +0.62), hence in both, the information that possession carried about goal difference was information already contained in team strength. The result scale is different: in cross-validated ordinal models, adding possession reduced the log loss for win, draw or loss in 2022 (1.004 to 0.941) but not in 2026 (0.882 to 0.898), for reasons examined in Section 3.4.

Table 3. Nested models of match goal difference. Possession coefficient (goals per percentage point) with 95% CI; robust standard errors.

Model	**2022 (n = 64)**	**2026 (n = 104)**	**Pooled (n = 168)**
M0: Elo gap + host	– adj. R^2 0.19; CV-RMSE 1.73	– adj. R^2 0.42; CV-RMSE 1.61	– adj. R^2 0.33; CV-RMSE 1.68
M1: + possession	−0.024 [−0.083, +0.034] adj. R^2 0.20; CV-RMSE 1.77 F test p = 0.24	+0.004 [−0.027, +0.036] adj. R^2 0.41; CV-RMSE 1.63 F test p = 0.80	−0.007 [−0.037, +0.023] adj. R^2 0.33; CV-RMSE 1.69 F test p = 0.58
M2: + ten tactical variables	+0.113 [−0.019, +0.245] adj. R^2 0.42; CV-RMSE 1.67 F test p = 0.002	+0.096* [+0.006, +0.186] adj. R^2 0.56; CV-RMSE 1.49 F test p < 0.001	+0.105** [+0.035, +0.176] adj. R^2 0.52; CV-RMSE 1.47 F test p < 0.001
M3: + shots, shots on target	+0.024 [−0.143, +0.190] adj. R^2 0.51; CV-RMSE 1.60 F test p = 0.006	+0.062 [−0.042, +0.165] adj. R^2 0.62; CV-RMSE 1.42 F test p < 0.001	+0.054 [−0.022, +0.130] adj. R^2 0.59; CV-RMSE 1.38 F test p < 0.001
M4: + xG (2026 only)	–	+0.068 [−0.024, +0.159] adj. R^2 0.68; CV-RMSE 1.29 F test p < 0.001	–

*p<0.05, **p<0.01, ***p<0.001. F tests compare each model with the one above it. The pooled models allow the Elo coefficient to differ by tournament. CV-RMSE: repeated ten-fold cross-validated root mean squared error of goal difference.

The pooled interaction model tests the cross-tournament question directly. Possession × tournament was +0.031 (95% CI −0.036 to +0.097; p = 0.37) and Elo × tournament was +0.061 (95% CI −0.201 to +0.323; p = 0.65); jointly, F = 0.85, p = 0.43. On the goal-difference scale, therefore, there is no evidence that either possession or team strength was valued differently in the two tournaments. This is absence of evidence rather than evidence of equality: the confidence interval for possession × tournament extends to +0.097, larger than the

whole unadjusted 2026 coefficient. Separately, each 100 Elo points was worth +0.413 goals in 2022 and +0.501 in 2026, and the residual standard deviations were similar (1.70 and 1.59 goals; Levene p = 0.87). Any difference between the tournaments is therefore more likely to reside in how goal margins translated into wins, draws and losses, which Sections 3.4 and 3.5 examine.

The distinction between the raw and adjusted relationships suggests that possession's apparent value is strongly dependent on the competitive composition of the matches in which it is observed. In 2026, possession was positively correlated with goal difference (r = 0.41), but this association nearly disappeared after accounting for Elo differences. Because possession was similarly correlated with Elo in both tournaments (r = 0.61 and 0.62), the contrast between the tournaments cannot be attributed simply to stronger teams having greater possession in one tournament than the other. Instead, it reflects differences in how team strength, possession and match outcomes coincided. The higher proportion of possession-dominant winners in 2026 is therefore compatible with possession having little additional predictive value once pre-match strength is known. This distinction is consequential: a statistic may discriminate successfully between winners and losers in an unadjusted comparison while providing almost no improvement over a model using information available before the match. Possession's descriptive association with winning and its incremental predictive value are therefore not interchangeable measures of its importance.

3.3 Total, shared and unique information in possession

Possession's coefficient changes with what else is held constant, and the changes are informative (Figure 2). Once the ten tactical variables enter (M2), possession becomes positive: +0.113 (95% CI −0.019 to +0.245; p = 0.09) in 2022, +0.096 (95% CI +0.006 to +0.186; p = 0.04) in 2026 and +0.105 (95% CI +0.035 to +0.176; p = 0.004) pooled. Among teams with the same pass completion and the same phase-of-play profile, the one that held more of the ball tended to do better. When shots and shots on target are added (M3) the coefficient falls to +0.024 (p = 0.78), +0.062 (p = 0.24) and +0.054 (p = 0.16), consistent with the conditional benefit of possession operating through the chances it produces. M3 estimates a residual association with chance creation held fixed, not possession's total association, and a coefficient that disappears at that step does not show that possession carries no information; it shows where its association runs.

The reversal in the possession coefficient between M1 and M2, followed by its attenuation in M3, is particularly informative. Possession is not statistically isolated from tactical organization: teams that retain the ball also tend to complete more passes, progress through different phases of play and create different attacking opportunities. Conditioning on these correlated characteristics changes the comparison being made. M1 compares teams of similar pre-tournament strength, whereas M2 compares teams that are additionally similar in their observed tactical profiles. The positive possession coefficient in M2 therefore describes a narrower conditional association,

not evidence that possession becomes causally beneficial when tactical characteristics are held constant. Introducing shots and shots on target in M3 changes the comparison again by conditioning on variables closely related to scoring production. The resulting attenuation is consistent with possession and chance creation sharing predictive information, but does not identify the proportion of possession's effect mediated through chances. The sequence of models thus demonstrates that the apparent contribution of possession depends substantially on which components of football performance are treated as explanatory variables and which are treated as intermediate outcomes.

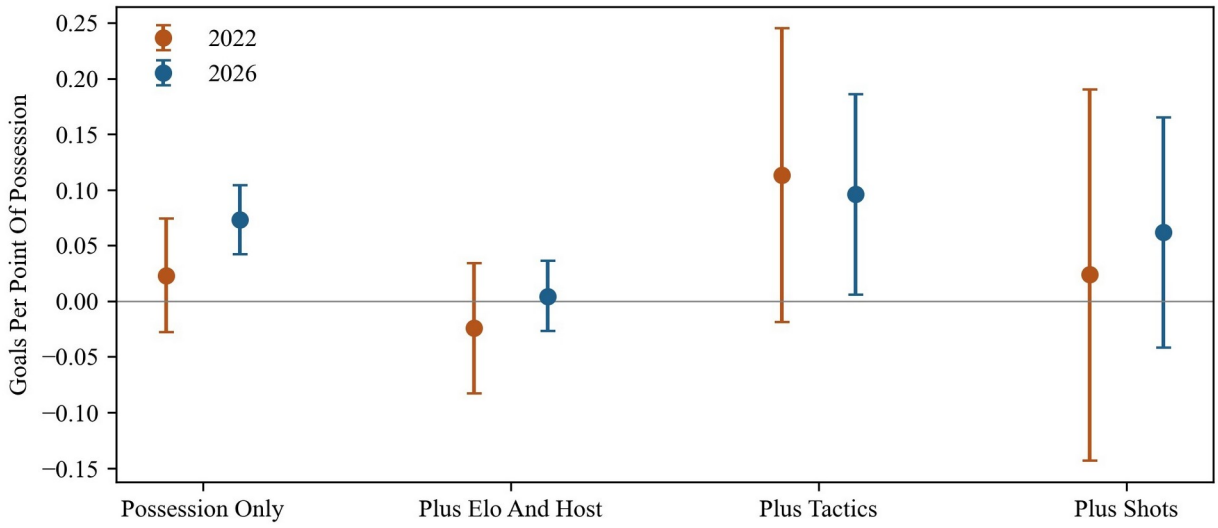


Figure 2. Possession coefficient (goal difference per percentage point of possession, with 95% confidence interval) as quality, tactical variables and shots are added, by tournament.

Residualised possession asks whether teams that held more of the ball than expected for their quality and style did better. Out-of-fold, quality and style explained 91% of the variance in possession. The unexpected remainder was not associated with goal difference in 2022 (+0.045; 95% CI −0.066 to +0.157; p = 0.42), in 2026 (+0.064; 95% CI −0.019 to +0.146; p = 0.13) or pooled (+0.044; 95% CI −0.027 to +0.116; p = 0.22). Adding passing volume, which explains possession almost mechanically (out-of-fold R^2 0.97), left the conclusion unchanged (−0.086 (p = 0.40) and +0.110 (p = 0.26)). The residual possession component showed no statistically distinguishable association with goal difference in these specifications. This supports the interpretation that possession contains substantial information shared with team strength and tactical organization, but does not establish that possession could not affect outcomes if deliberately changed.

Commonality analysis quantifies that reading (Figure 3). With four predictor sets (possession; Elo gap and host; progression, the

line-break and final-third rates; chances, shots and shots on target), possession's unique contribution to R^2 was 0.002 in 2022, 0.0005 in 2026 and 0.001 pooled, against total R^2 of 0.50, 0.56 and 0.51. Its shared contribution was 0.023, 0.192 and 0.110. In 2026, possession alone explained 19% of the variance in goal difference, almost all of it shared with strength, progression and chances. General dominance weights rank possession last of the four sets in 2022 and pooled, and third in 2026 (0.057, against 0.208 for quality and 0.256 for chances).

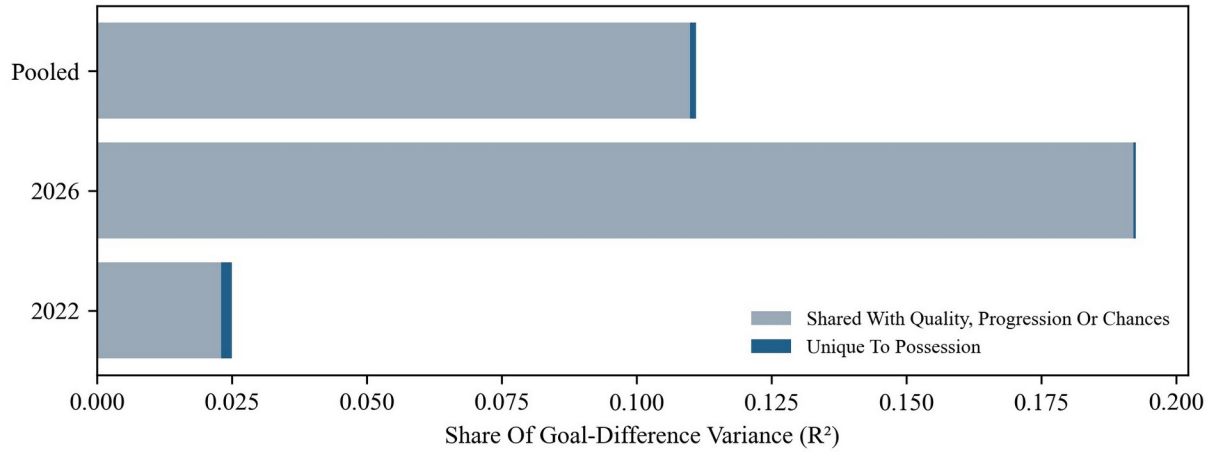


Figure 3. Commonality analysis of goal difference: variance explained by possession that is unique to it and variance it shares with team strength, progression and chance creation.

A principal-component analysis of fourteen team-level style measures retained 3 components with eigenvalues above one, explaining 71% of the variance. After varimax rotation the first component is a sustained-control dimension on which possession loads 0.88, with passes (0.87) and pass completion (0.91) loading positively and attacking transition (−0.92) and long-ball share (−0.85) negatively. The second is a chance-production dimension (directness 0.92, shots 0.82, final-third rate 0.77); the third is defensive posture (PPDA 0.77, low block 0.56, forced turnovers −0.55). Possession is therefore one observable face of a broader territorial-control construct. Given the Elo gap, the chance-production component predicted goal difference (+0.343; $p < 0.001$ pooled) while the control component did not (−0.146; $p = 0.14$). Adding possession itself to the three components gave a positive coefficient in every sample (2022: $b = +0.201$, $p = 0.02$; 2026: $b = +0.215$, $p < 0.001$), the same pattern as M2: once style is held fixed, the possession a team held beyond what its style implies was associated with better results.

Elastic-net regression resulted in the same answer from a prediction standpoint. Possession was retained in 48% of bootstrap samples in 2022 and 89% in 2026, with standardised coefficients of 0.00 and 0.24, but dropping it changed cross-validated RMSE by

at most 0.008 goals (1.544 versus 1.537 in 2022; 1.427 versus 1.435 in 2026). Elastic net was preferred to the lasso because it can retain correlated predictors together; tree-based models were not used.

3.4 The 2022 adjusted coefficient

An earlier version of this manuscript reported an Elo-adjusted possession coefficient of −0.049 (95% CI −0.091 to −0.007; p = 0.02) for 2022 and then described possession as having no association. That coefficient originated from the 62-match sample. With all 64 matches and the same end-of-year Elo it is −0.035 (95% CI −0.082 to +0.013; p = 0.15), and with pre-tournament Elo and the host indicator it is −0.024 (95% CI −0.083 to +0.034; p = 0.42). The estimate is sensitive to individual matches: leave-one-out estimates range from −0.045 to −0.019, and the most influential observation is Spain versus Costa Rica (DFBETA 1.18), a 7–0 win with a large possession advantage. On the goal-difference scale, then, the 2022 coefficient is negative in sign but not distinguishable from zero.

On the result scale the picture is sharper, and it should not be described as zero. In an ordinal model of loss, draw or win, Elo-adjusted possession in 2022 was −0.087 (95% CI −0.142 to −0.031; p = 0.002) log-odds per percentage point; among decided matches the logistic coefficient was −0.090 (95% CI −0.160 to −0.019; p = 0.01). In 2026 both were near zero (−0.002 (p = 0.92) and +0.012 (p = 0.77)), and the pooled ordinal possession × tournament interaction was +0.088 (p = 0.01). In 2022, holding team strength fixed, the team with more of the ball was less likely to win. Goal difference dilutes this because favourites' wins carried large margins (2.3 goals on average in 2022) while 10 of the 11 2022 upsets were by a single goal. Section 3.5 identifies the matches responsible.

Three further explanations were examined. Negative confounding is expected: the bivariate coefficient is +0.023 (p = 0.37), and because possession correlates +0.61 with the Elo gap, adjustment removes the part of possession that marks strength and can reverse the sign of a small remainder. Game state may also contribute. Unconditionally, teams that scored first held slightly more of the ball (51.7% against 48.3% in 2022), which reflects their greater strength. Holding the Elo gap fixed, scoring first was associated with 2.2 fewer points of possession in 2022 (p = 0.12) and 2.1 fewer in 2026 (p = 0.06), the direction expected if leading teams cede the ball, and adjusting for which team scored first moves the 2022 goal-difference coefficient to −0.010 (p = 0.73). Because whole-match possession cannot separate possession held while level from possession held while trailing, reverse causation cannot be ruled out. Finally, model instability and multiple testing are real concerns for a single coefficient in 64 matches, which is why the upset analysis in the next sections examines the upset matches directly rather than relying on this coefficient.

3.5 How strength was converted into results: the upsets

The two tournaments did not differ in the distribution of Elo gaps (Kolmogorov-Smirnov p = 0.82) or in residual variance of goal difference, and the share of matches decided by

at most one goal was similar (0.61 and 0.57), as was the knockout share (0.25 and 0.31). They may have differed in how often the favourite won, although the evidence is not conclusive. The Elo favourite won 35 of 49 decided matches in 2022 (71.4%) and 69 of 80 in 2026 (86.2%; z = 2.07, p = 0.04; Fisher exact p = 0.06). Against the Elo formula's own expectations, favourites underperformed by 0.069 points per match in 2022 and overperformed by 0.038 in 2026, and the Brier score of Elo's predictions was 0.187 against 0.108. In an ordinal model the Elo × tournament interaction was +0.260 (p = 0.08), in the direction of weaker conversion in 2022 but not conclusive. Figure 4 shows the pattern across gap sizes; in 2022 underdogs won 24% of matches with a 200 to 399 point gap, against 8% in 2026.

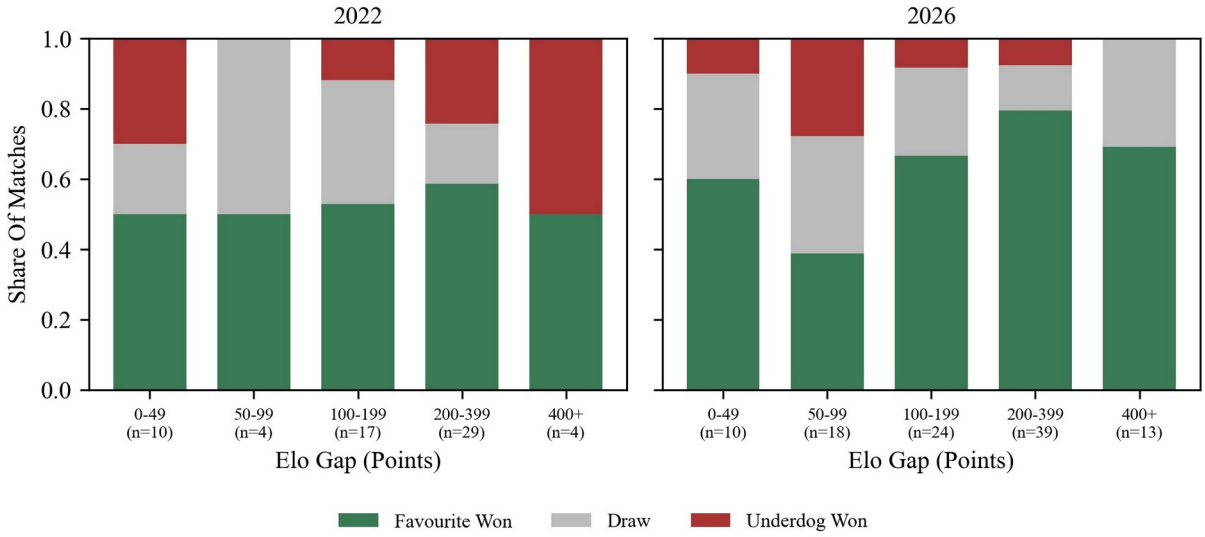


Figure 4. Results from the Elo favourite's point of view, by size of the pre-tournament Elo gap. Penalty shoot-outs are counted as draws.

With a 50-point threshold for a clear favourite, 2022 produced 11 upsets, 13 draws and 30 confirming results, and 2026 produced 10, 21 and 63. The upset rate was 20.4% against 10.6% (Fisher p = 0.14); with no threshold it was 21.9% against 10.6% (p = 0.07), and with a 100-point threshold 22.0% against 6.6% (p = 0.01). The excess of upsets in 2022 is consistent across definitions though conventionally significant only at the stricter threshold.

Table 4 compares upsets with confirming results. In 2022 the underdogs that won had a possession deficit of 36.7 points, against 10.5 for underdogs that lost, and they spent more time than their opponents in counter-attack, attacking-transition, long-ball and low-block phases, forced more turnovers, and usually scored first (64% of upsets). These differences survive false-discovery-rate adjustment. In 2026 none of the possession or transition differences appear: underdogs that won had

nearly the same possession deficit as those that lost, and what distinguished them was that they matched or exceeded the favourite in shots on target (+0.8 against −3.8) and xG (+0.32 against −1.34), and scored more than their xG implied. Scoring first was common to upsets in both tournaments.

Table 4. Upset versus Elo-confirming matches (favourite by at least 50 Elo points). Mean underdog-minus-favourite difference in upsets vs confirming results.

Underdog minus favourite	2022 (11 upsets, 30 confirming)	2026 (10 upsets, 63 confirming)
Elo gap (favourite minus underdog, points)	+284 vs +244	+127 vs +251**
Possession gap (points)	−36.7 vs −10.5**	−6.4 vs −14.2
Counter-attack share (points)	+1.82 vs −0.73**	+0.30 vs +0.22
Attacking-transition share (points)	+11.5 vs +1.0**	+2.9 vs +3.6
Long-ball share (points)	+5.0 vs +1.1*	+2.1 vs +2.2
Low-block share (points)	+17.5 vs +5.0*	+7.3 vs +9.7
Forced turnovers (SD units)	+1.46 vs −0.15**	+0.13 vs −0.27
Directness (shots per 100 passes)	+0.79 vs −0.51	+0.53 vs −0.59
Shots	−5.7 vs −5.1	−1.1 vs −6.7
Shots on target	−1.27 vs −3.47	+0.80 vs −3.81**
Expected goals	–	+0.32 vs −1.34**
Goals minus xG	–	+1.28 vs −0.82***
Underdog scored first (share)	+0.64 vs +0.07***	+0.90 vs +0.17***
Minute of first goal	+41 vs +34	+41 vs +30
Set-piece phase share (points)	+1.8 vs +0.2	−0.2 vs +1.0
Red cards	+0.18 vs +0.00*	+0.10 vs +0.08
Penalty goals	−0.09 vs −0.10	+0.00 vs −0.13

Asterisks mark Mann-Whitney tests significant after Benjamini-Hochberg adjustment across all comparisons: *q<0.05, **q<0.01, ***q<0.001. xG is published for 2026 only.

Upsets occurred at different gaps in the two tournaments: the mean Elo gap was 284 points for 2022 upsets against 244 for 2022 confirming results, but 127 against 251 in 2026. Because underdogs facing substantially stronger opponents tend to have less possession regardless of their tactical approach, the comparisons were repeated after adjusting for the Elo gap and allowing its association with the outcomes to vary between tournaments. The 2022 upset-versus-confirming difference in the possession gap remained −25.4 (95% CI −38.0 to −12.9; $p < 0.001$) points, with a tournament interaction of +26.1 (95% CI +2.0 to +50.2; $p = 0.03$); for attacking-transition share the 2022 difference was +10.4 ($p < 0.001$) and the interaction −9.5 ($p = 0.04$).

Pooled logistic models of upset versus confirming result, with the Elo gap, tournament

and their interaction, point the same way. The possession gap predicted an upset in 2022 (−0.063; p = 0.005 per point) with a positive interaction for 2026 (+0.063; p = 0.04); the corresponding interactions were −0.610 (p = 0.04) for counter-attack share and −0.230 (p = 0.004) for attacking transition, while the effect of scoring first did not differ (−0.002; p = 1.00). This is the pattern expected if 2022 contained successful low-possession, transition-based performances by underdogs and 2026 largely did not. With 11 and 10 upsets respectively, these estimates rest on few matches and should be read as a consistent descriptive pattern rather than as precise effects.

The upset analysis suggests that the difference between tournaments was not primarily a general change in the frequency of counterattacking football, but a difference in the characteristics of the performances that produced unexpected victories. In 2022, successful underdogs had an average possession deficit of 36.7 percentage points, compared with 10.5 points among underdogs in Elo-confirming results. Their greater counterattacking and attacking-transition shares indicate that the low-possession victories were associated with a recognizable tactical profile rather than merely an absence of ball control. Importantly, these teams did not generally overcome their opponents through greater shot volume. Their success was instead concentrated in converting relatively limited attacking opportunities into favorable results. In 2026, successful underdogs were more distinguishable by their ability to match or exceed favorites in shots on target and expected goals. The distinction is therefore between different observed routes to an upset: successful outcomes despite limited attacking volume in 2022, and stronger measured chance production among successful underdogs in 2026. Because successful outcomes were used to define the upset groups, these comparisons cannot establish that adopting either tactical profile would increase an underdog's probability of winning. They do, however, explain why the same possession statistic acquired different associations with match results across the tournaments.

3.6 Strategy: frequency versus payoff

The four profiles (Table 5, Figure 5) show that the strategies occurred about equally often in the two tournaments ($\chi^2 = 0.68$, df = 3, p = 0.88); what changed was their payoff. Productive control, high possession with above-median shots, won 42% of the time in 2022 and 63% in 2026. Efficient transition, low possession with above-median shots, won 55% in 2022 and 25% in 2026, and teams that conceded the ball and created little won 34% and 16%. Of the 14 underdog wins in 2022, 10 came from performances with low possession and below-median shots: these teams won by converting few chances, not by creating many. The median splits are for interpretation only; the continuous models remain primary. Using xG instead of shots to define chance production in 2026 gave win rates of 70%, 26%, 44%, 9% for high possession and high xG, high possession and low xG, low possession and high xG, and low possession and low xG).

Comparing how often each style occurred (match averages across both teams, false-

discovery-rate adjusted), 2022 had more counter-attack time (1.75% against 1.46%, q = 0.006), whereas 2026 had more low-block time (23.2% against 19.1%) and more receptions in the final third per 100 passes. Attacking transition, directness and long-ball share did not differ. Among Elo underdogs specifically, counter-attack, attacking-transition and directness levels did not differ between tournaments, while 2026 underdogs spent more time in a low block (28.2% against 23.0%, q = 0.018). Underdogs therefore did not play more directly in 2022 on average; what differed was which underdogs won. Forced turnovers averaged 71 per team in 2022 and 42 in 2026, a difference too large to be tactical; it almost certainly reflects a change in FIFA's definition, so cross-tournament comparisons of that variable's level are not interpreted. Per standard deviation, counter-attack share predicted goal difference similarly in both tournaments (+0.294 ($p = 0.02$) and +0.352 ($p < 0.001$); difference $p = 0.72$), while the line-break rate predicted it in 2022 only (difference −0.345; $p = 0.03$). The coefficients describe the value of a style per unit; the distributions describe how often it was used. Neither shows that 2022 favoured rapid attack in general. The difference lies in who used transition play successfully: the underdogs that won in Table 4.

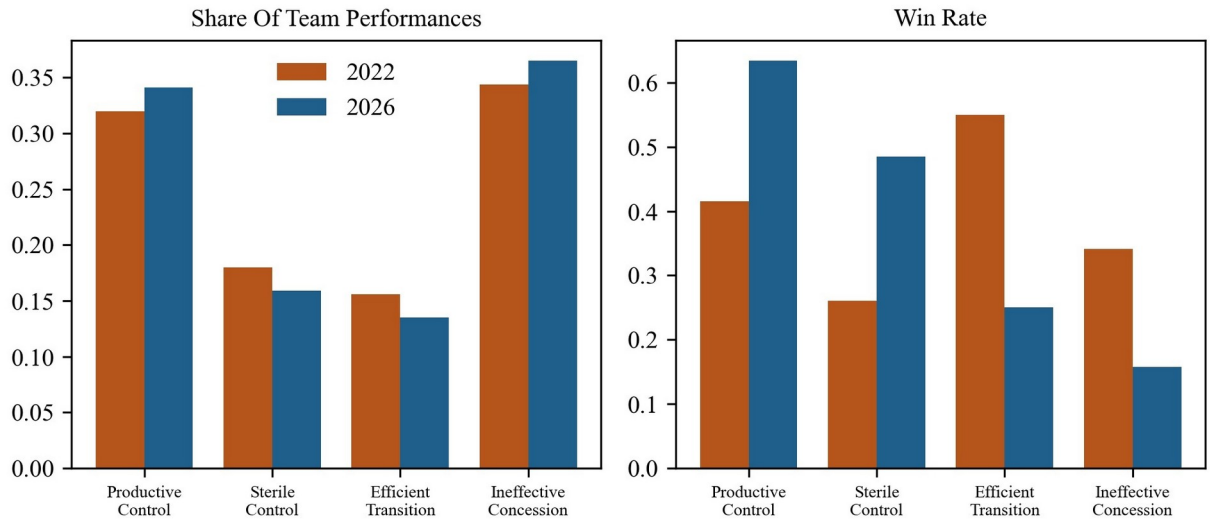


Figure 5. Four strategic profiles of team performances: share of all team performances (left) and win rate (right), by tournament. High possession means above 50%; high chance production means shots above the tournament median.

Possession × strategy interactions were tested for ten moderators in each tournament, with a test of whether each interaction differed between tournaments, plus xG per shot in 2026: 31 tests in all (Table 6). Of these, 4 were nominally significant, and 1 remained significant after either Benjamini-Hochberg or Holm adjustment. The omnibus tests that all ten interactions are zero gave $p = 0.39$ in 2022 and $p = 0.08$ in 2026. The surviving interaction is with having scored first in 2026 (+0.096; 95% CI +0.061 to +0.131; $p < 0.001$): possession was associated with better results for teams that took the lead. Because the first

goal partly determines later possession, this is more plausibly a game-state effect than evidence that possession pays only when ahead. The nominal interaction with chance quality in 2026 (+0.355; p = 0.03) did not survive adjustment (p = 0.28 adjusted).

Table 5. Four strategic profiles of team performances (team-match level).

Profile	Year	n	Share	Win rate	Mean GD	Mean Elo gap	Underdog wins
Productive control	2022	41	0.32	0.41	+0.44	+166	1
	2026	71	0.34	0.63	+1.28	+184	3
Sterile control	2022	23	0.18	0.26	−0.35	+85	0
	2026	33	0.16	0.48	+0.39	+76	2
Efficient transition	2022	20	0.16	0.55	+0.85	−56	3
	2026	28	0.14	0.25	−0.36	−59	3
Ineffective concession	2022	44	0.34	0.34	−0.61	−174	10
	2026	76	0.36	0.16	−1.24	−183	3

High possession: above 50%. High chance production: shots above the tournament median. Elo gap is the team's rating minus its opponent's. Underdog wins: wins by the lower-rated team.

Table 6. Possession × strategy interactions.

Moderator	2022	2026	Difference, 2026 minus 2022
Directness	−0.019 [−0.062, +0.023] p 0.373; q 0.729; Holm 1.000	−0.009 [−0.025, +0.007] p 0.263; q 0.703; Holm 1.000	+0.011 [−0.033, +0.055] p 0.625; q 0.799; Holm 1.000
Line-break rate	−0.005 [−0.018, +0.008] p 0.489; q 0.758; Holm 1.000	−0.002 [−0.007, +0.002] p 0.318; q 0.703; Holm 1.000	+0.002 [−0.012, +0.017] p 0.740; q 0.882; Holm 1.000
Final-third rate	−0.003 [−0.009, +0.004] p 0.418; q 0.729; Holm 1.000	−0.001 [−0.005, +0.004] p 0.812; q 0.900; Holm 1.000	+0.002 [−0.005, +0.010] p 0.547; q 0.771; Holm 1.000
Counter-attack share	+0.000 [−0.061, +0.061] p 0.989; q 0.989; Holm 1.000	−0.016 [−0.035, +0.002] p 0.086; q 0.536; Holm 1.000	−0.015 [−0.077, +0.048] p 0.645; q 0.799; Holm 1.000
Attacking-transition share	−0.003 [−0.012, +0.007] p 0.589; q 0.793; Holm 1.000	+0.003 [−0.002, +0.008] p 0.287; q 0.703; Holm 1.000	+0.006 [−0.004, +0.017] p 0.236; q 0.703; Holm 1.000
Long-ball share	+0.007 [−0.011, +0.026] p 0.435; q 0.729; Holm 1.000	−0.004 [−0.014, +0.005] p 0.382; q 0.729; Holm 1.000	−0.011 [−0.032, +0.010] p 0.294; q 0.703; Holm 1.000
High-press share	+0.007 [−0.006, +0.020] p 0.305; q 0.703; Holm 1.000	+0.007 [+0.000, +0.014] p 0.038; q 0.292; Holm 1.000	+0.000 [−0.015, +0.015] p 0.984; q 0.989; Holm 1.000
Low-block share	+0.000 [−0.004, +0.005] p 0.865; q 0.925; Holm 1.000	−0.002 [−0.004, +0.000] p 0.126; q 0.650; Holm 1.000	−0.002 [−0.008, +0.003] p 0.447; q 0.729; Holm 1.000
Forced turnovers	−0.021 [−0.052, +0.010] p 0.182; q 0.703; Holm 1.000	+0.003 [−0.016, +0.021] p 0.776; q 0.891; Holm 1.000	+0.024 [−0.011, +0.060] p 0.182; q 0.703; Holm 1.000
Scored first	+0.016 [−0.036, +0.068] p 0.540; q 0.771; Holm 1.000	+0.096 [+0.061, +0.131] p <0.001; q <0.001; Holm <0.001	+0.079 [+0.016, +0.141] p 0.013; q 0.209; Holm 0.405
Chance quality (xG per shot)	–	+0.355 [+0.040, +0.670] p 0.027; q 0.282; Holm 0.792	–

Each model includes the Elo gap, host, possession and the moderator. Interaction terms are differences of team-level centred products. coefficient [95% CI], unadjusted p, Benjamini-Hochberg q and Holm-adjusted p.

3.7 Possession efficiency

How productively a team used the ball predicted results whereas the amount did not. Holding the Elo gap and possession fixed, shots on target per 100 passes predicted goal difference in 2022 (+1.250; 95% CI +0.692 to +1.809; $p < 0.001$) and in 2026 (+1.169; 95% CI +0.724 to +1.614; $p < 0.001$), as did directness (+0.351 ($p = 0.009$) and +0.294 ($p = 0.01$)) and, in 2026, xG per 100 passes (+4.185; $p < 0.001$). In each of these models possession itself was near zero. Because ratio variables can behave poorly when denominators vary, the equivalent non-ratio models with possession, chance production and their interaction were also fitted; the possession × shots interaction was −0.004 ($p = 0.37$) in 2022 and −0.004 ($p = 0.07$) in 2026, and possession × xG in 2026 was +0.015 ($p = 0.43$). Possession did not become more valuable when coupled to higher chance production; there was no evidence that possession's value depended on how many chances a team created. Among the progression rates, the line-break rate predicted goal difference in 2022 (+0.158; $p = 0.003$) but not in 2026 (+0.032; $p = 0.49$), and the final-third rate in neither (+0.062 ($p = 0.07$) and −0.009 ($p = 0.65$)).

3.8 Path decomposition

An earlier version of this manuscript stated that the indirect effect through xG was ~ the entire unadjusted possession coefficient. That was incorrect: the indirect and direct effects sum to the quality-adjusted total, not the unadjusted association. With the corrected sample and pre-tournament Elo, in 2026 each percentage point of possession was associated with +0.0294 xG holding quality constant, and each unit of xG with +0.917 goals holding possession and quality constant. The indirect path is +0.0269 (match-level bootstrap 95% CI +0.0092 to +0.0467) and the direct path −0.023 (95% CI −0.053 to +0.007; $p = 0.13$); together they equal the quality-adjusted total of +0.004 (95% CI −0.027 to +0.036; $p = 0.79$), far below the unadjusted +0.073. A positive indirect path offset by a negative direct path is consistent with possession helping a team create chances while, at a given level of chances, more possession accompanies worse results, as it would if possession is partly accumulated while chasing a match. Using shots on target, which is available for both tournaments, the indirect path was +0.0147 (95% CI −0.0054 to +0.0348) in 2022 and +0.0271 (+0.0127 to +0.0442) in 2026.

A path model for 2026 makes the assumptions explicit: team strength and score state (which team scored first) to possession; possession, strength and score state to progression (final-third rate); progression, possession, strength and score state to xG; and all of these to goal difference. Estimated links: strength to possession +3.374 ($p < 0.001$); scoring first to possession −2.055 ($p = 0.06$); possession to progression +0.454 ($p < 0.001$); progression to xG +0.033 ($p < 0.001$); scoring first to xG +0.299 ($p = 0.002$); xG to goal difference +1.053 ($p < 0.001$); possession directly to goal difference −0.008 ($p = 0.59$). All variables are whole-match aggregates measured at the same time, hence the ordering of these links is assumed, not observed. This is an observational path decomposition consistent with possession supporting progression and chance creation; it does not show that possession causes chances,

or that the whole of its association runs through xG.

3.9 Where possession carries information: territory, chances and goals

Table 7. Ablation: explained variance and prediction error with and without possession, 2026 matches with a clear favourite (n = 94), underdog minus favourite. Baseline: Elo gap and directness (shots per 100 completed passes).

Outcome	R^2 without / with (gain)	CV-RMSE without / with	Possession per 10 points [95% CI]	All 2026 matches, R^2 gain	2022, R^2 gain
Completed line breaks	0.269 / 0.806 (0.537***)	34.30 / 17.91	+15.4 [+13.2, +17.6]	0.380***	0.405***
Final-third receptions	0.242 / 0.766 (0.524***)	98.92 / 56.31	+43.0 [+35.8, +50.2]	0.352***	0.313***
Expected goals	0.428 / 0.568 (0.140***)	0.90 / 0.79	+0.232 [+0.148, +0.317]	0.079***	–
Forced turnovers	0.103 / 0.118 (0.015)	6.91 / 6.93	−0.5 [−1.2, +0.3]	0.014	0.330***
Goal difference	0.214 / 0.227 (0.013)	1.57 / 1.58	+0.104 [−0.092, +0.301]	0.006	0.001

Asterisks mark the nested F test for adding possession: *p<0.05, **p<0.01, ***p<0.001. CV-RMSE: repeated ten-fold cross-validated root mean squared error, in each outcome's own units. Robustness columns: all 104 2026 matches and all 64 2022 matches, team A minus team B, with the host indicator added to the baseline (uncentred R^2, since these models have no intercept). xG is published for 2026 only. Forced-turnover counts are defined differently in the two tournaments (Section 2.3).

The ablation separates three kinds of outcome (Table 7, Figure 6). For territory, possession carried a significant amount of information beyond strength and directness. Adding it raised the explained variance of completed line breaks from 0.27 to 0.81 and of final-third receptions from 0.24 to 0.77, and it reduced cross-validated prediction error by 48% and 43%. Ten more points of possession were associated with 15.4 more completed line breaks and 43.0 more final-third receptions. For chance quality the contribution was smaller but clear: the R^2 of xG increased from 0.43 to 0.57 ($p < 0.001$) and its prediction error decreased from 0.895 to 0.786, with +0.23 xG per ten points. For goals it was negligible: the gain in R^2 was 0.013 ($p = 0.23$) and cross-validated error was 1.575 without possession and 1.583 with it. Forced turnovers, which a team wins without the ball, gained 0.015 ($p = 0.23$). The same ordering held in all 104 matches of 2026 and in 2022 (Table 7, last two columns), where possession added 0.41 to the R^2 of line breaks and 0.001 to that of goal difference. Possession is therefore highly informative about where a team plays and moderately informative about the quality of the chances it creates, while adding almost nothing to the prediction of the final score once strength and directness are known.

The progressive attenuation of possession's incremental explanatory contribution is consistent with the increasing selectivity of

successive attacking outcomes. Completed line breaks and final-third receptions measure the frequency with which teams advance the ball into potentially productive positions. Additional possession can generate more such events simply by providing more opportunities to circulate and progress the ball. Expected goals impose a more restrictive criterion: territorial advancement must result in shots with measurable scoring probability. Realized goals introduce further variation because even comparable expected-goal totals can produce different scores through finishing outcomes, goalkeeper performance and the probabilistic nature of individual chances. Consequently, possession may strongly predict the volume of intermediate attacking events without strongly predicting the final score. The much larger R^2 gains for progression than for goals are consistent with this interpretation, although the gains should not be treated as directly comparable measures of causal influence because the outcomes differ in their variability, measurement and dependence on possession. The central finding is that possession's predictive information becomes progressively less distinctive as the outcome moves from territorial activity toward realized scoring.

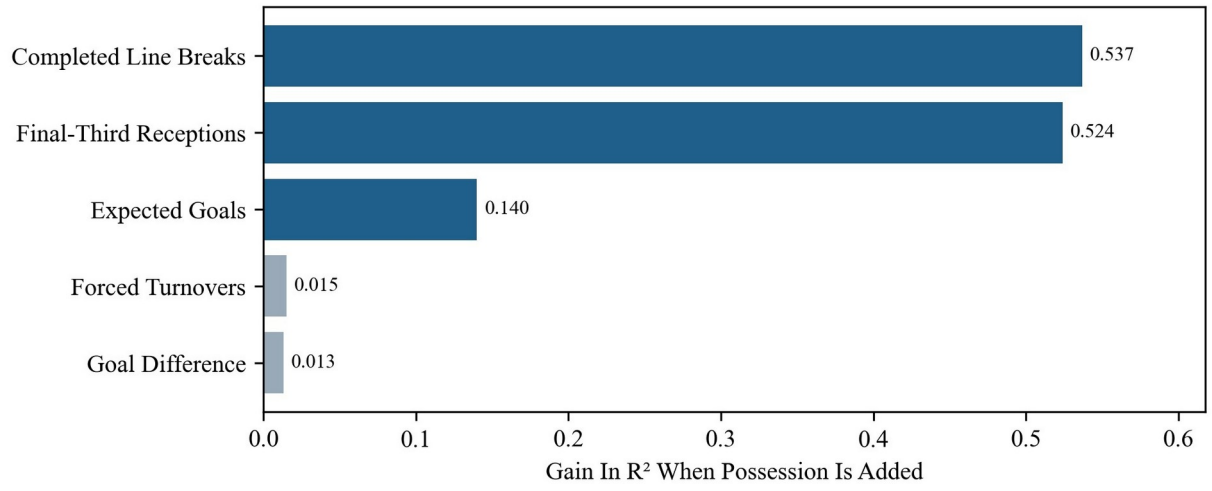


Figure 6. Ablation: gain in explained variance (R^2) when possession is added to a model of each underdog-minus-favourite outcome containing the Elo gap and directness, 2026 matches with a clear favourite (Elo gap of at least 50 points; 94 matches). Dark bars: nested F test $p < 0.05$; light bars: not significant.

Territorial progression and possession share substantial predictive information about xG, but the targeted ablation does not establish how much of possession's association with xG operates through territorial progression. Line breaks and final-third receptions alone raised the R^2 of xG from 0.428 to 0.544; adding possession as well raised it to 0.570 (nested F test $p = 0.02$), with a possession coefficient of +0.210 (95% CI +0.012 to +0.409; $p = 0.04$) xG per ten points, against +0.232 without the territorial measures. Line breaks (−0.0010; $p = 0.84$) and final-third receptions (+0.0010; $p = 0.58$) were not individually significant in the

full model, and possession's variance inflation factor there was 6.4, hence the three measures largely carry the same information and their separate coefficients are unstable. Held-out prediction error changed from 0.817 to 0.802 when possession was added to the territorial model. It would be incorrect either to conclude that territory fully explains possession's link with chance quality or that possession has a firmly identified effect of its own on it.

Possession has a defensive side as well as an attacking one, since time with the ball is time the opponent cannot attack. Because possession is zero-sum, a regression of chances conceded on possession is simply the mirror image of a regression of chances created and cannot separate the two. Instead, each team's xG was written as its rate per unit of possession multiplied by its possession time, and the association of possession with the xG difference was divided into three parts, from the point of view of the team with more of the ball: extra attacking time at its own rate, time denied to the opponent at the opponent's rate, and any change in the rate itself, estimated with the Elo gap and host held fixed. This uses every match of the tournament. In 2026, ten points of possession were worth +0.140 (95% CI +0.127 to +0.153; $p < 0.001$) xG through attacking time and +0.107 (95% CI +0.093 to +0.122; $p < 0.001$) through denial, so about 43% of possession's volume benefit was defensive; the rate change was −0.082 (95% CI −0.189 to +0.012; $p = 0.09$), and the net +0.164 is close to the DML estimate below. For goals the pattern differed in one respect. The attacking-time and denial parts were +0.166 and +0.123, but the rate change was −0.231 (95% CI −0.373 to −0.089; $p = 0.001$): the more of the ball a team had, the fewer goals it scored per unit of possession, leaving a net of +0.058 (95% CI −0.085 to +0.201; $p = 0.42$). In 2022 the offset was larger. The team with less of the ball scored more per unit of possession than its opponent (3.78 against 2.62 goals per full match of possession), so denial (+0.168) outweighed attacking time (+0.118), and with a rate change of −0.511 ($p < 0.001$) the net was −0.226 (95% CI −0.456 to +0.035; $p = 0.09$). This is the same pattern seen on the result scale in Section 3.4 and in the low-possession upsets of Section 3.5. This decomposition is an accounting identity evaluated with whole-match rates, not a causal model, and a falling rate may reflect diminishing returns, sterile circulation or possession accumulated while trailing. It does, however, show where possession's value for goals goes: the extra chances it brings are offset by lower conversion per unit of possession.

The negative rate component deserves particular caution because it may reflect several mechanisms that cannot be separated using whole-match totals. One possibility is tactical: additional possession may occur in situations where defenses are organized and attacking opportunities are comparatively difficult to convert. Another is selection by match state: teams that fall behind may accumulate possession while attempting to recover the score, whereas teams protecting a lead may deliberately accept less possession and attack selectively. A third possibility is mathematical. Because the scoring rate is calculated by dividing goals by possession time, differences

in the denominator can generate an inverse association even without a corresponding causal decline in attacking effectiveness. The negative rate component therefore should not be interpreted as evidence that increasing possession necessarily makes a team less efficient. Its importance is instead that the positive volume components do not translate proportionately into realized goals in the observed matches. Distinguishing tactical inefficiency from game-state selection and mathematical coupling would require event-level measures of attacking exposure and scoring production. The present decomposition establishes how the observed association is distributed among its accounting components, but not which mechanism generated those components.

Table 8. Double machine learning estimates, 2026 matches with a clear favourite (n = 94): change in each underdog-minus-favourite outcome per ten points of possession difference, with 95% CI.

Adjustment set	Outcome	Ridge	Random forest	OLS
Elo gap, home status	Completed line breaks	+14.9 [+12.9, +16.9]***	+15.0 [+12.9, +17.0]***	+14.9***
	Final-third receptions	+38.9 [+31.7, +46.0]***	+40.6 [+32.9, +48.3]***	+39.1***
	Expected goals	+0.15 [+0.06, +0.24]**	+0.13 [+0.03, +0.23]*	+0.15**
	Goal difference	+0.02 [−0.16, +0.21]	−0.03 [−0.20, +0.14]	+0.02
+ line breaks, final-third receptions	Expected goals	+0.00 [−0.20, +0.20]	−0.07 [−0.28, +0.15]	−0.01
	Goal difference	−0.04 [−0.37, +0.29]	−0.12 [−0.46, +0.21]	−0.05
+ directness	Expected goals	+0.22 [+0.04, +0.40]*	+0.09 [−0.10, +0.27]	+0.22*
	Goal difference	+0.25 [−0.09, +0.58]	+0.10 [−0.22, +0.42]	+0.24

Five-fold cross-fitting repeated 50 times; median estimate with a variance that includes the spread across splits. *p<0.05, **p<0.01, ***p<0.001. The first block is the primary analysis. The other two condition on variables measured during the match, which may lie on the path from possession to chances, and are sensitivity analyses rather than estimates of possession's total effect. OLS: ordinary least squares with the same covariates.

Double machine learning asks whether these associations depend on a linear adjustment for strength. With only pre-match covariates (Table 8, Figure 7), ten more points of possession were associated with 14.9 (ridge) and 15.0 (random forest) more completed line breaks, 38.9 and 40.6 more final-third receptions, and +0.148 (p = 0.002) and +0.131 (p = 0.01) xG. The goal-difference estimates were +0.023 (95% CI −0.16 to +0.21) and −0.031 (−0.20 to +0.14). The two learners agree with each other and with ordinary least squares, hence the pattern is not an artefact of linear control for strength. The conclusions were unchanged when the host indicator replaced FIFA's nominal home designation (goal difference +0.001, ridge), when all 104 matches of 2026 were analysed in the team-A-minus-team-B design (line breaks +14.8, xG +0.146, goal difference +0.024; p = 0.76), and

in 2022 (line breaks +14.5; goal difference −0.118; p = 0.41).

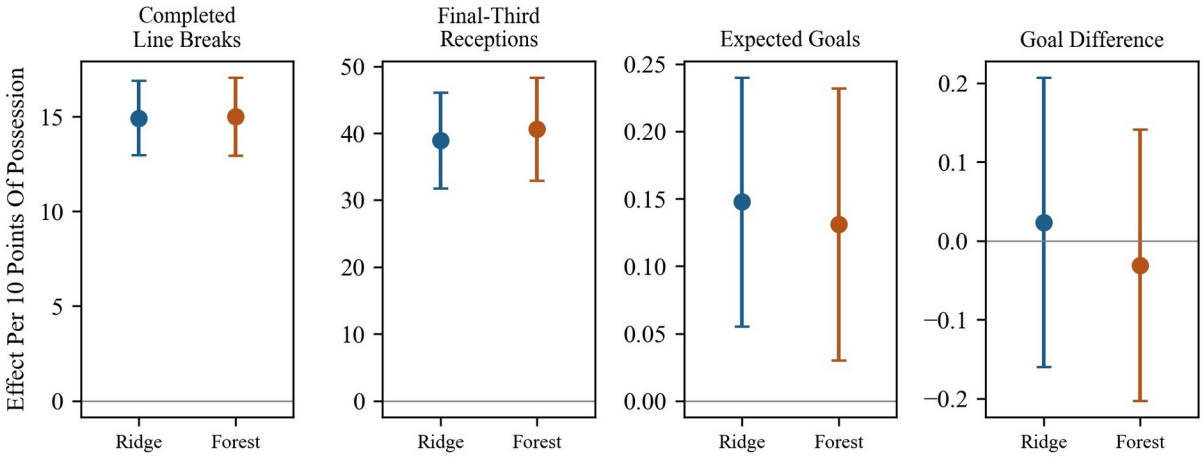


Figure 7. Double machine learning estimates: change in each underdog-minus-favourite outcome per ten points of possession difference, adjusted for the Elo gap and home status with ridge regression and random forest nuisance models, with 95% confidence intervals, 2026 matches with a clear favourite (94 matches). Panels have different scales.

The conditional specifications show how much the decomposition depends on what is held fixed. Adding line breaks and final-third receptions to the adjustment set moved the xG estimate to +0.000 with ridge (p = 1.00) and −0.066 with random forest (p = 0.54). Adding directness as well gave +0.221 with ridge (p = 0.02) and +0.089 with random forest (p = 0.34). For goal difference the estimates were −0.039 with ridge (p = 0.82) and −0.121 with random forest (p = 0.48) with territory, and +0.247 with ridge (p = 0.15) and +0.102 with random forest (p = 0.53) with territory and directness. These variables are measured during the match and may be consequences of possession, causes of it, or both, hence conditioning on them can remove a genuine pathway or create a spurious one. The pre-match specification estimates possession's total adjusted association; the conditional specifications describe how that association overlaps with territory and directness, and neither is a causal effect. The sensitivity of the estimates to these choices is itself a finding: how possession's association is divided depends on the tactical variables held fixed.

3.10 Sensitivity analyses

The 2022 and 2026 conclusions are robust to the strength measure. With each team's rating before each match, Elo-adjusted possession was −0.021 (p = 0.48) in 2022 and +0.000 (p = 0.99) in 2026; with end-of-year ratings, −0.034 (p = 0.19) and +0.009 (p = 0.54). Adding nonlinear Elo terms gave −0.031 (p = 0.31) and +0.006 (p = 0.71); the nonlinear terms themselves mattered in 2022 (p = 0.001), reflecting large-gap upsets such as Cameroon 1-0 Brazil and Argentina 1-2 Saudi Arabia.

Using 90-minute goal difference gave −0.024 (p = 0.42) and +0.002 (p = 0.88).

Matched and weighted comparisons at team-performance level, pairing each higher-possession team with a lower-possession team from another match of the same tournament and stage with an Elo gap within 25 points, resulted in a goal-difference advantage for the higher-possession team of −0.79 (bootstrap 95% CI −1.41 to −0.09; 34 pairs) in 2022 and +0.46 (+0.00 to +0.91; 74 pairs) in 2026; inverse-probability weighting resulted in −1.14 and +0.67. These comparisons do not depend on a linear model and point the same way as the result-scale models: among teams of similar relative strength, more possession was associated with worse outcomes in 2022 and somewhat better ones in 2026. The samples are too small for richer matching, so these are corroborating rather than primary analyses. Requiring the matched teams also to agree on whether they scored first resulted in −0.56 (−1.41 to +0.41) in 2022 and +0.57 (+0.09 to +1.05) in 2026.

3.11 Limitations

This is an observational study of two tournaments. Teams choose how to play, and stronger teams choose differently, hence every coefficient describes an association under statistical control rather than the effect of a team changing its approach. Whole-match possession cannot separate possession held while level from possession accumulated after falling behind; the published reports contain no time-resolved possession, so the game-state analyses above cannot rule out reverse causation, and this limitation constrains the causal language of every conclusion. Expected goals are available for 2026 only, so the xG decomposition rests on one tournament, and the variables available for both tournaments were used for every cross-tournament comparison. Some FIFA metrics, most clearly forced turnovers, appear to have been defined differently in the two tournaments. Set-piece goals could not be identified consistently and were represented only by FIFA's set-piece phase share and penalty goals. Finally, the upset analysis rests on 21 upsets, and the cross-tournament differences it reveals, though consistent across several measures, need confirmation in further tournaments.

The ablation and double machine learning analyses share these limits. Whole-match possession is measured after many goals and tactical decisions have already occurred: an underdog that scores first and then defends deep will finish with less of the ball, so its possession deficit can be a consequence of its success rather than a cause. Neither method resolves this feedback with aggregate totals, and the double machine learning estimates would be causal only under no unmeasured confounding, correct temporal ordering and overlap, none of which can be defended here; they are adjusted associations. Pre-match strength and home or host status are a limited adjustment set, and adding variables measured during the match changes the estimates, which shows that the decomposition depends on the assumed causal ordering. The attacking and defensive parts of possession were separated only by an accounting identity using whole-match rates. Teams appear in several matches, so observations are not fully independent, and

the many models reported here were not adjusted for multiplicity beyond the analyses that state it. A more credible causal analysis would use time-resolved possession and events, segment matches by score state, and relate possession in one period to chances and goals in the next; this is a requirement for future research with event-level data.

3.12 Perspectives: Beyond *Total Football*—toward *Sublime Football*

3.12.1 *The historical primacy of possession*

The evolution of modern football has been characterized by increasingly sophisticated strategies for acquiring, retaining, recovering, and denying possession. The Dutch development of Total Football, associated particularly with Rinus Michels and Johan Cruyff, emphasized positional interchange, spatial control, coordinated movement, and collective responsibility for attacking and defending. Cruyff's subsequent influence on positional play, particularly at Barcelona, helped establish the philosophical foundations for later possession-oriented systems, including the tiki-taka style associated with Pep Guardiola's Barcelona. High pressing, counterpressing, and tactical structures employing formations such as the 4-3-3 and 3-4-3 further illustrate the importance attached to recovering possession and restricting an opponent's ability to use the ball. These tactical developments were not concerned exclusively with accumulating possession.

Total Football and its successors also emphasized creating space, disrupting defensive organization, generating numerical superiority, and producing scoring opportunities. Nevertheless, possession became a prominent operational objective and a widely reported indicator of tactical dominance. The underlying assumption was that sustained control of the ball would facilitate attacking production while simultaneously restricting the opponent's opportunities.

The present study provides an opportunity to reconsider this assumption. If possession is so central to football's tactical evolution, why does it contribute so little additional explanatory information about realized goal difference? The apparent contradiction may arise because possession has been evaluated simultaneously as a prerequisite for attacking play, a tactical resource, and a statistical predictor of match outcomes. These are related but fundamentally different properties.

3.12.2 *Possession as a necessary resource rather than an ordinary predictor*

A fundamental distinction exists between possession as an enabling condition for attacking production and possession as a statistical predictor of goals. Conventional regression, ablation analysis, and double machine learning estimate whether differences in possession explain additional variation in match outcomes after accounting for other variables. They do not establish whether possession is necessary for the production of those outcomes.

In ordinary attacking play, some degree of ball control is generally required to create and complete a scoring opportunity, even if that control lasts only a fraction of a second. A

successful counterattack may involve very little possession, whereas an extended passing sequence may produce no meaningful opportunity. Strict necessity must nevertheless be qualified: own goals, deflections, and the operational definitions used to record possession mean that positive measured possession is not an absolute prerequisite for every goal. The more defensible proposition is that effective attacking access to the ball is ordinarily an enabling condition for goals generated by a team's own attacking actions.

The distinction between necessity and statistical association is recognized in Necessary Condition Analysis (NCA), which investigates whether a minimum level of a condition is required to make a specified outcome feasible, rather than whether increasing that condition raises the average outcome (11,12). A necessary condition may impose a boundary on possible outcomes without exhibiting a substantial regression coefficient. Thus, the statistical insignificance of an explanatory variable does not establish its structural irrelevance to the process generating the outcome.

However, this principle does not establish that any particular percentage of match possession is necessary for scoring. The minimum effective attacking opportunity required to score may be extremely small, and possession percentages conflate productive attacks, defensive circulation, and tactical responses to the score. Formal NCA using whole-match possession percentages could therefore fail to capture the relevant enabling condition. A more appropriate analysis would examine individual attacking sequences, effective ball control, and the minimum opportunity required to generate particular levels of attacking production.

This distinction suggests that possession should be evaluated through three separate scientific questions: [1] *Necessity*: What minimum attacking opportunity or effective ball control is required for a specified scoring outcome to become feasible? [2] *Productivity*: How efficiently is available attacking possession transformed into territorial progression, expected goals, and realized goals? and [3] *Incremental prediction*: Does possession explain additional variation in match outcomes after accounting for team strength and other observed characteristics?

The present manuscript's ablation and double machine learning analyses principally address the third question. Its time-and-rate decomposition begins to illuminate the second. The first remains a conceptual and empirical question requiring an appropriately designed investigation.

The proposed distinction between possession as an enabling condition, a productive tactical resource, and a statistical predictor also connects with established research on causality and football performance. Causal inference requires distinguishing observed associations from the mechanisms generating them (13). Earlier football research examined the effectiveness of alternative playing strategies (14) and the relationships between passing sequences, shots, and goals (15), while subsequent work demonstrated how possession strategies vary with match score and situational

conditions (16). More recently, action-valuation methods have sought to quantify the contribution of individual football actions to competitive outcomes rather than relying exclusively on aggregate performance statistics (17). These analytical developments complement the historical evolution of football tactics from Total Football toward increasingly sophisticated systems of positional control, attacking organization, and defensive pressure (18). Together, they provide a foundation for examining possession not simply as time on the ball, but as a resource whose value depends on the opportunities and competitive advantages it generates.

3.12.3 *The progression from territorial dominance to scoring outcomes*

The manuscript's findings reveal a pronounced attenuation of possession's incremental predictive contribution across successive stages of attacking production. Among the 94 matches from the 2026 World Cup involving a clear favorite, adding possession to models containing team-strength differences and attacking directness increased explained variance by 0.537 for completed line breaks and 0.524 for final-third receptions. The corresponding increase was 0.140 for expected goals and only 0.013 for realized goal difference. Thus, possession was highly informative about territorial progression, moderately informative about chance production, and contributed little additional predictive information about the final score.

The double machine learning analysis supported the same hierarchy. With pre-match team strength and home status included in the adjustment set, a ten-percentage-point possession difference was associated with approximately 15 additional completed line breaks, 39–41 additional final-third receptions, and 0.13–0.15 additional expected goals. The corresponding estimates for goal difference were close to zero and statistically indistinguishable from zero. The agreement between ridge regression and random forest adjustment indicates that the principal pattern was not merely an artifact of assuming a linear relationship between team strength and the outcomes.

These findings should not be interpreted as evidence that possession is unimportant or dispensable. Rather, they demonstrate that possession's incremental predictive value depends on the stage of attacking production being examined. Ball control facilitates territorial access and the creation of attacking opportunities, but progression, opportunity quality, defensive resistance, finishing, and chance variation intervene before those opportunities become realized goals.

Furthermore, the modest incremental contribution of possession to goal difference must be distinguished from its total strategic contribution. Team strength, tactical organization, territorial progression, and possession are interrelated. Statistical adjustment may remove variation that is shared among these variables, while conditioning on in-match progression measures may also remove pathways through which possession operates. Consequently, a small adjusted coefficient does not establish a small total causal effect. The present observational

analyses identify predictive relationships rather than independently validated causal mechanisms.

The central question is therefore not simply whether possession predicts victory, but why its strong association with territorial advancement becomes progressively weaker as the outcome approaches realized goals.

3.12.4 *Possession productivity: From ball retention to goal production*

One explanation lies in the distinction between possession volume and possession productivity. A useful conceptual representation is; G = T × R, where G = goals produced through attacking possession; T = effective attacking possession time and R = goal production rate per unit of effective attacking possession.

This multiplicative relationship distinguishes the amount of an available attacking resource from the rate at which that resource is converted into goals. A team may double its effective attacking possession without increasing its scoring output if its production rate falls proportionately. Conversely, a team may score more with substantially less possession if its attacking sequences generate and convert opportunities more efficiently.

This relationship is an accounting identity when the production rate is defined as goals divided by effective possession time; it does not establish that possession time causally determines goals. Furthermore, because possession time appears in the denominator of the rate, a negative relationship between possession volume and measured productivity may arise partly through mathematical coupling. Evidence of genuinely diminishing tactical returns therefore requires analyses that distinguish arithmetic effects from differences in opportunity quality, attacking strategy, opponent strength, and match state.

The manuscript's decomposition provides a particularly informative extension of this framework. In the 2026 tournament, a ten-percentage-point possession difference was associated with an expected-goal contribution of approximately +0.140 through additional attacking time and +0.107 through denying the opponent possession time. The latter represented approximately 43% of the combined positive volume component. A negative rate component of approximately −0.082 offset part of these benefits, leaving a net expected-goal difference of approximately +0.164.

The corresponding decomposition for realized goals was more striking. Additional attacking time contributed approximately +0.166 goals and defensive denial approximately +0.123 goals, but the rate component was −0.231, leaving a net association of only +0.058 goals. In the 2022 tournament, the negative rate component was even larger, consistent with the observed success of some low-possession underdogs and the importance of counterattacking and attacking transitions.

These estimates demonstrate why the offensive and defensive volume benefits associated with possession need not translate into equivalent scoring advantages. However, the decomposition is based on whole-match

accounting relationships rather than experimentally identified causal effects. Lower scoring rates among higher-possession teams may reflect diminishing attacking returns, but they may also arise because trailing teams accumulate possession against organized defenses, leading teams deliberately concede possession, or different tactical strategies produce different distributions of opportunities.

The findings therefore support a distinction between possession accumulation and possession productivity, without establishing that increased possession itself causes productivity to decline.

3.12.5 *The dual strategic value of possession*

Possession has two inseparable strategic dimensions. It provides the possessing team with an opportunity to attack, while simultaneously denying the opponent the opportunity to construct attacks through its own possession. The first dimension is offensive production; the second is defensive denial.

This duality helps explain why football teams have historically invested so heavily in positional control, high pressing, and rapid ball recovery. Regaining possession does more than initiate another attack. It also interrupts the opponent's attacking process and changes the distribution of available attacking time.

Yet the defensive value of possession is not equivalent to preventing goals. A team may dominate possession while remaining vulnerable to counterattacks, defensive transitions, or a small number of high-quality opportunities. Conversely, a team that deliberately concedes possession may reduce the opponent's attacking productivity through compact defensive organization while increasing its own scoring efficiency through rapid transitions.

The tactical value of possession must therefore be assessed through its combined offensive and defensive consequences. A strategy that increases possession but reduces the team's attacking productivity or exposes it to dangerous transitions may provide less competitive advantage than a strategy involving shorter, more productive possession sequences.

This suggests that possession should be regarded as a resource with potentially competing returns: the benefit of retaining the ball, the opportunity cost of prolonged circulation, the defensive benefit of denying possession, and the vulnerability created when possession is lost. The balance among these returns is likely to depend on team capabilities, opponent organization, and match circumstances.

3.12.6 *Sublime Football: A proposed evolution beyond Total Football*

The foregoing distinctions motivate a proposed conceptual evolution that may be termed Sublime Football.

Total Football helped establish coordinated movement, positional interchange, collective pressing, and spatial control as foundations of modern tactical organization. Subsequent possession-oriented systems refined the

acquisition, retention, circulation, and recovery of the ball. Sublime Football would make the efficient transformation of possession into competitive advantage the explicit objective of tactical optimization.

Its central principle is that possession should neither be maximized as an end in itself nor minimized in pursuit of attacking directness. Instead, possession should be acquired, retained, accelerated, or relinquished according to its expected contribution to attacking production and defensive restriction.

Total Football sought to control the ball; Sublime Football seeks to control what the ball accomplishes.

This proposed framework would evaluate tactical performance across several interconnected dimensions: possession volume, territorial penetration, attacking directness, opportunity quality, scoring productivity, defensive denial, and transition vulnerability. A team retaining possession for prolonged periods would be considered tactically effective only insofar as that possession contributes to favorable attacking and defensive outcomes. Similarly, a team employing rapid counterattacks would be evaluated not by its low possession percentage, but by the competitive advantage it generates from the opportunities available.

Sublime Football does not imply that earlier tactical philosophies neglected scoring efficiency or purposeful possession. Nor does the present study demonstrate that a distinct optimal tactical system has been discovered. The proposed contribution is instead a quantitative reframing: possession becomes an enabling resource whose effectiveness is evaluated through its conversion into competitive advantage, rather than a standalone measure of tactical superiority.

3.12.7 *The possession–productivity frontier*

An extension of Sublime Football is the hypothesis that teams operate along a *possession–productivity frontier*, describing the combinations of possession volume and attacking efficiency associated with the greatest attainable expected scoring advantage under particular tactical conditions.

At low possession volumes, additional effective attacking time may offer substantial opportunities for progression and chance creation. At higher volumes, the marginal benefit of additional possession may diminish if opponents become increasingly organized, attacks become more predictable, or prolonged circulation replaces more productive attacking transitions. Alternatively, sustained possession may remain highly productive for teams possessing exceptional technical ability, coordinated positional movement, and mechanisms for disrupting defensive structures.

The relevant optimum would therefore depend on the team's capabilities and the opponent's defensive and attacking characteristics. It would also depend on the score and remaining match time. A team protecting a lead may rationally prioritize possession's defensive denial value, whereas a trailing team may need to sacrifice retention efficiency to increase the

frequency and immediacy of scoring opportunities.

Conceptually, the objective would be to maximize:
Expected competitive advantage = Expected attacking production − Expected opponent attacking production.

Possession would be one controllable tactical resource affecting both terms. The objective is not equivalent to maximizing expected goals alone, because attacking decisions may also alter defensive exposure and the opponent's subsequent opportunities.

The present findings motivate this frontier hypothesis but do not establish its existence, shape, or optimal operating points. Such claims would require additional empirical work, particularly analyses capable of separating tactical choices from the match conditions that induce them.

3.12.8 *A research agenda for testing Sublime Football*

A direct empirical test would require time-resolved possession sequences rather than whole-match possession percentages alone. Individual sequences could be classified by duration, starting field position, territorial progression, directness, opponent defensive organization, score state, and resulting attacking or defensive outcomes.

Three related investigations would be particularly informative.

First, a necessity analysis could determine whether measurable minimum levels of effective attacking opportunity are associated with specified levels of chance production. This would test whether the relevant constraint is possession duration, number of attacking sequences, territorial access, or another operational measure of attacking opportunity. The analysis should distinguish genuine necessity relationships from constraints that arise mechanically from variable definitions.

Second, possession productivity could be estimated as a function of tactical context. Rather than treating goals per minute of possession as an independent explanatory variable, a suitable statistical model could treat effective possession time as exposure while modeling expected goals or scoring-event rates as functions of directness, progression, opponent organization, and match state. Nonlinear relationships could then be examined to determine whether additional possession is associated with diminishing, constant, or increasing attacking returns.

Third, the proposed possession–productivity frontier could be investigated by comparing alternative tactical strategies within similar competitive circumstances. Possession during one period could be related to subsequent chances and goals, while accounting for preceding score state, team strength, and tactical context. Such temporal separation would help reduce the reverse-causality problem created when teams change their possession strategies in response to the score. A stronger causal investigation would additionally require credible identification of

tactical changes that are not simply responses to unobserved match conditions.

Together, these investigations could determine whether an identifiable tactical operating range exists in which possession's marginal attacking and defensive benefits are jointly maximized. They could also test whether the optimal range differs between teams, opponents, and match states, rather than assuming a universally desirable possession percentage.

3.12.9 *Broader implications*

The principal implication of the present study is that possession's structural importance, tactical productivity, and incremental predictive significance must be distinguished. The fact that possession adds little explanatory information about realized goal difference after adjustment for team strength does not mean that possession is unimportant to the production of goals. Conversely, the fact that effective ball control ordinarily enables attacking opportunities does not mean that accumulating more possession necessarily improves scoring outcomes.

The manuscript's empirical contribution is to demonstrate how possession's incremental explanatory information attenuates from territorial progression to chance quality and finally to realized goal difference, while also identifying substantial defensive denial and opposing productivity components. Its broader conceptual contribution is to suggest that possession should be studied not simply as a percentage associated with winning, but as a strategic resource whose value depends on its offensive and defensive deployment.

This interpretation also helps reconcile the apparent differences between the 2022 and 2026 World Cups. The success of low-possession underdogs in 2022 and the stronger descriptive association between possession and winning in 2026 need not represent contradictory principles of football. They may instead reflect different combinations of team strength, tactical strategies, opportunity production, defensive denial, and scoring productivity. The available whole-match analyses cannot establish which mechanisms caused these differences, but they demonstrate why possession percentage alone is insufficient to characterize tactical effectiveness.

Sublime Football offers a conceptual direction for extending this research. It proposes that the next advance in football analytics should move beyond measuring control of possession toward evaluating the competitive returns generated by that control. The essential distinction is between possessing the ball and converting possession into favorable outcomes. In this framework, the ultimate tactical objective is neither possession dominance nor attacking directness in isolation, but the efficient coordination of attacking production, defensive denial, and scoring advantage.

The evolution from Total Football to Sublime Football is therefore a progression from controlling possession to optimizing its consequences.

4 Conclusion

Possession appeared inconsequential at the 2022 World Cup but decisive in 2026. Neither interpretation withstands adjustment for team

strength. In both tournaments, possession was similarly correlated with team strength and contributed negligible additional predictive information about goal difference once strength was accounted for. Its unique contribution to explained variance was at most 0.2%, whereas its shared contribution with team strength, territorial progression, and chance creation was substantial. The pooled analysis found no statistically significant difference between tournaments in the associations of either possession or team strength with goal difference, although a moderate difference could not be excluded.

On the win-draw-loss scale the tournaments did differ. In 2022, among teams of similar strength, the team with more of the ball was less likely to win, and in 2026 it was not. That pattern coincides with the 2022 upsets, which had a recognisable statistical profile even after allowing for the Elo gap: the underdogs that won had much less of the ball than underdogs that lost and spent more of their time in counter-attacking and attacking-transition phases. In 2026 the underdogs that won were distinguished instead by matching their opponents' shots on target and expected goals. These are whole-match descriptions; they do not show the order in which events happened within matches.

Examining where possession carries information is more revealing than asking whether possession itself is useful. Among 2026 matches with a clear favourite, possession explained an additional 54% of the variance in completed line breaks, 14% in expected goals (xG), and only 1.3% in goal difference after accounting for team strength and directness. Double machine learning, allowing more flexible adjustment for team strength, reproduced this hierarchy. Possession also has a defensive dimension: time spent controlling the ball denies the opponent attacking opportunities. However, the additional chances associated with greater possession were largely offset by lower goal production per unit of possession. Thus, possession can strongly characterize how a match unfolds while contributing little additional information about its final outcome.

Possession is best understood neither as a decisive asset nor as an independently established cause of winning, but as an observable indicator of a coordinated strategic state. Its meaning plausibly depends on whether it is coupled to progression and chance production, accumulated while chasing a match, or conceded to an opponent built for transition, and the mix of those situations in a tournament is a likely reason why possession can look valuable in one tournament and not in the next.

These findings also reveal a fundamental distinction between possession's necessity, productivity and statistical significance. Effective ball control is ordinarily an enabling condition for generating attacking opportunities, even when the period of possession is extremely brief. However, the necessity of an enabling resource does not imply that increasing its quantity will proportionately increase the outcome it enables. Conversely, the negligible incremental predictive contribution of possession to

realized goal difference does not establish its structural irrelevance to attacking production. The distinction between a necessary condition and a statistically significant predictor therefore helps resolve the apparent paradox at the center of this study: possession can be fundamental to the generation of scoring opportunities while explaining little additional variation in realized goals. The relevant question is not simply how much possession a team accumulates, but how effectively it transforms that possession into attacking production and defensive advantage.

This distinction provides the conceptual foundation for Sublime Football, proposed here as an evolution beyond possession-centered tactical philosophies. Building on Total Football's emphasis on positional control, interchange and collective recovery of possession, Sublime Football proposes that possession should be operationalized to maximize its combined offensive and defensive returns rather than maximized as an end in itself. The associated possession–productivity frontier is a testable hypothesis that the most advantageous balance between possession volume, attacking efficiency and defensive denial depends on team capabilities, opponent organization and match state. Establishing such a frontier will require time-resolved data and methods capable of separating tactical decisions from their consequences. Total Football sought to control the ball; Sublime Football seeks to control what the ball accomplishes. The broader contribution of this study is therefore to redirect football analytics from possession as a measure of dominance toward possession as a strategic resource whose value is determined by its conversion into territorial progression, attacking opportunities, defensive restriction and ultimately scoring advantage.

Acknowledgements
The author thanks the reviewer for constructive comments and suggestions.

Abbreviations
CI: confidence interval, CV: cross-validated, DML: double machine learning, Elo: World Football Elo rating, GD: goal difference, OLS: ordinary least squares, PPDA: opponent passes per defensive pressure, RMSE: root mean squared error, xG: expected goals.

Data availability
The cleaned match-level dataset, data dictionary, extraction code, analysis code and the source reports' URLs are deposited at Zenodo (https://doi.org/10.5281/zenodo.23302561), and every number in this paper is regenerated by the deposited scripts. In addition, data and code are also deposited in https://github.com/bobrsultan/What-Possession-Is-Worth-Evidence-from-Two-World-Cups

The deposited package contains: data/all_links.txt (the URLs of the 168 FIFA reports) and fetch_sources.py, which downloads the FIFA reports, Elo ratings, openfootball results and Wikipedia line-ups, since these third-party files are not redistributed; data/fifa_reports.csv

(statistics extracted from the reports); data/matches.csv (the cleaned one-row-per-match dataset) and data/DATA_DICTIONARY.md; build_dataset.py (extracts team statistics from the reports); build_matches.py (merges results, Elo and red cards and checks every score against FIFA); analysis_revision.py (every analysis, table value and figure in this paper, with a fixed random seed); analysis_r3.py (the ablation, time-and-rate decomposition and double machine learning analyses of Section 3.9 and Figures 6 and 7, run after analysis_revision.py); and results/results.json, results/log.txt and results/log_r3.txt (the complete output). Running python3 fetch_sources.py, python3 build_dataset.py, python3 build_matches.py, python3 analysis_revision.py and python3 analysis_r3.py in order reproduces every number; the last two alone reproduce them from the deposited data/matches.csv. The core model specification is:

```
import pandas as pd, statsmodels.formula.api as smf
M = pd.read_csv("data/matches.csv")
X = pd.DataFrame({"gd": M.goals_a - M.goals_b, "elo": (M.elo_pre_a - M.elo_pre_b) / 100,
                  "poss": 100 * M.poss_a_raw / (M.poss_a_raw + M.poss_b_raw) - 50, "year":
M.year})
HOSTS = {2022: {"Qatar"}, 2026: {"USA", "Mexico", "Canada"}}
X["host"] = [int(a in HOSTS[y]) - int(b in HOSTS[y]) for a, b, y in zip(M.team_a, M.team_b,
M.year)]
X["t26"] = (X.year == 2026).astype(int)
m1 = smf.ols("gd ~ 0 + elo + host + poss", X[X.year == 2026]).fit(cov_type="HC3")
pooled = smf.ols("gd ~ 0 + elo + poss + host + elo:t26 + poss:t26", X).fit(cov_type="HC3")
```

References

1. Collet, C. (2013). The possession game? A comparative analysis of ball retention and team success in European and international football, 2007–2010. *Journal of Sports Sciences, 31*(2), 123–136. https://doi.org/10.1080/02640414.2012.727455

2. Fédération Internationale de Football Association. (n.d.). *The relationship between possession and winning.* FIFA Training Centre. https://www.fifatrainingcentre.com/

3. Degrenne, O., & Carling, C. (2024). Comparison of goalscoring patterns between the 2018 and 2022 FIFA World Cups. *Frontiers in Sports and Active Living, 6*, Article 1394621. https://doi.org/10.3389/fspor.2024.1394621

4. Fédération Internationale de Football Association. (2022). *Post-match summary reports: FIFA World Cup Qatar 2022.* FIFA Training Centre.

https://www.fifatrainingcentre.com/en/fwc2022/post-match-summaries/post-match-summary-reports.php

5. Fédération Internationale de Football Association. (2026). *Match report hub: FIFA World Cup 2026.* FIFA Training Centre. https://www.fifatrainingcentre.com/en/fifa-world-cup-2026/match-report-hub.php

6. openfootball. (2026). *World Cup match data in JSON: 2022 and 2026* [Data set]. GitHub. https://github.com/openfootball/worldcup.json

7. Wikipedia contributors. (2026). *2022 and 2026 FIFA World Cup group and knockout stage articles: Line-ups.* Wikipedia. Retrieved October 10, 2026, from https://en.wikipedia.org/

8. World Football Elo Ratings. (2026). *Match results and ratings, 2021–2026* [Data set]. Retrieved October 10, 2026, from https://www.eloratings.net/

9. Seabold, S., & Perktold, J. (2010). statsmodels: Econometric and statistical modeling with Python. In *Proceedings of the 9th Python in Science Conference* (pp. 92–96). https://doi.org/10.25080/Majora-92bf1922-011

10. Chernozhukov, V., Chetverikov, D., Demirer, M., Duflo, E., Hansen, C., Newey, W., & Robins, J. (2018). Double/debiased machine learning for treatment and structural parameters. *The Econometrics Journal, 21*(1), C1–C68. https://doi.org/10.1111/ectj.12097

11. Dul, J. (2016). Necessary condition analysis (NCA): Logic and methodology of "necessary but not sufficient" causality. *Organizational Research Methods, 19*(1), 10–52. https://doi.org/10.1177/1094428115584005

12. Dul, J., van der Laan, E., & Kuik, R. (2020). A statistical significance test for necessary condition analysis. *Organizational Research Methods, 23*(2), 385–395. https://doi.org/10.1177/1094428118795272

13. Pearl, J. (2009). *Causality: Models, reasoning, and inference* (2nd ed.). Cambridge University Press.

14. Pollard, R., & Reep, C. (1997). Measuring the effectiveness of playing strategies at soccer. *Journal of the Royal Statistical Society: Series D (The Statistician), 46*(4), 541–550.

15. Hughes, M., & Franks, I. (2005). Analysis of passing sequences, shots and goals in soccer. *Journal of Sports Sciences, 23*(5), 509–514.

16. Lago-Peñas, C., & Dellal, A. (2010). Ball possession strategies in elite soccer according to the evolution of the match-score: The influence of situational variables. *Journal of Human Kinetics, 25*, 93–100.

17. Decroos, T., Bransen, L., Van Haaren, J., & Davis, J. (2019). Actions speak louder than goals: Valuing player actions in soccer. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining* (pp. 1851–1861). Association for Computing Machinery.

18. Wilson, J. (2008). *Inverting the pyramid: The history of football tactics*. Orion.