After the First Touch: What Decides the Second Ball at Set Pieces

Jaime Oriol Goicoechea

1. Introduction

Corners are a set piece that coaches typically prepare in detail: who takes the kick, where the ball is aimed and which runs open space for whom. However, the first touch does not always settle the play. A defender may head the cross clear, a teammate may flick it on or a shot may be blocked, after which the ball drops and both teams contest it again. For example, after Crystal Palace conceded from an Arsenal corner that they had defended at the first contact, their manager Oliver Glasner said: "We defended the first ball but we didn't defend the second ball, the rebound" [1]. Yet, how much of the value of a corner is decided in this second phase, and whether a team can plan for it, remains unclear.

A common assumption is that what follows the first touch is largely a matter of chance. Getting a flick-on to fall where a teammate is standing, with no defender closer, involves more moving parts than crossing the ball onto a head. The second phase therefore depends on more factors than the delivery, and fewer of those factors are under a coach's control. In a game as studied as football, however, that lack of control may also be an asset. What a coach cannot script an opponent cannot fully prepare for either, and the uncertainty may surprise a side that has prepared for the delivery.

The question came from watching Arsenal's corners through the 2025/26 season. A coach can name the routines, such as the dummy, the blocks or the run to the far post that opens the near one. Yet something else seemed to decide the outcome: players crowded the box, someone got a touch, the ball broke loose and it went in.[1] Arsenal scored 19 goals from corners and Tottenham Hotspur 18, own goals included, both above the previous Premier League record of 16 [2]. Of these goals, 8 and 12 respectively arrived after the first touch, which are the joint second-highest and the highest counts among the 96 sides of the five big leagues.

Arsenal won the title by seven points, whereas Tottenham finished 17th, two points above the relegation zone, and in margins that small a single goal can decide a season. The two scored more goals from corners than any other Premier League side, which may well have made the difference to each club's own measure of success: the title for one and survival for the other. Hence, a high count of goals after the first touch does not seem to belong to successful teams alone. Two clubs cannot establish a general pattern, but they do raise a question about the second ball itself rather than about one team.

---

[1] Arsenal's video of its set-piece goals of the season, published on its own channel on 19 May 2026; the corner at 4:08 shows it clearly: https://www.youtube.com/watch?v=Mwtve2XkOus&t=248s

Previous work offers little evidence on this second phase. Power et al. [3] followed set pieces past the first contact, whereas models built on player positions at corners stop at the delivery or at the first contact [4, 5, 6]. To the best of my knowledge, no published model values the second ball from the freeze-frame taken at the kick, which shows where the players in the box stand.

I therefore ask two questions:

1. How much of the value of a corner is decided after the first contact, and how much of that was already set before the delivery?

2. When a team wins the second ball, is it because the plan put a player where the ball was always going to drop, or because someone reacted faster on the day?

Before the data were released, I wrote down what I expected as three hypotheses, one on the first question and two on the second:

H1. Where the attackers stand at the kick predicts how the corner ends, beyond where and how the ball is struck.

H2. A model of who can reach the loose ball first beats the static picture of where the bodies stand.

H3. A second ball won by a player sent to wait where it drops is worth more than one won by whoever is nearby.

To answer them, the event chain of each corner is split at the first touch. A shot comes *after the first touch* when a defender cleared the ball or an attacker kept it alive before the shot, within seconds of the first contact. The shots and goals on either side of the split are counted first, and a model then tests whether the freeze-frame taken at the kick predicts anything beyond what the delivery alone predicts. Together they give two main results:

1. Of the goals that corners produce, 41.1% come after the first touch (Figure 1). Per shot, the value sits in the attacking touch that keeps the ball alive: a shot after a flick-on converts 12.2 percentage points (pp) more often than a direct shot and one after a prolongation 9.4 pp more, while a shot after a clearance does not differ from a direct shot.

2. The positions at the kick predict where the ball is aimed but, once the delivery is known, add nothing measurable to how the corner ends, whereas the same method finds a clear signal in who wins the first contact. H1 and H3 are not supported, and H2 cannot be tested on thirty matches.

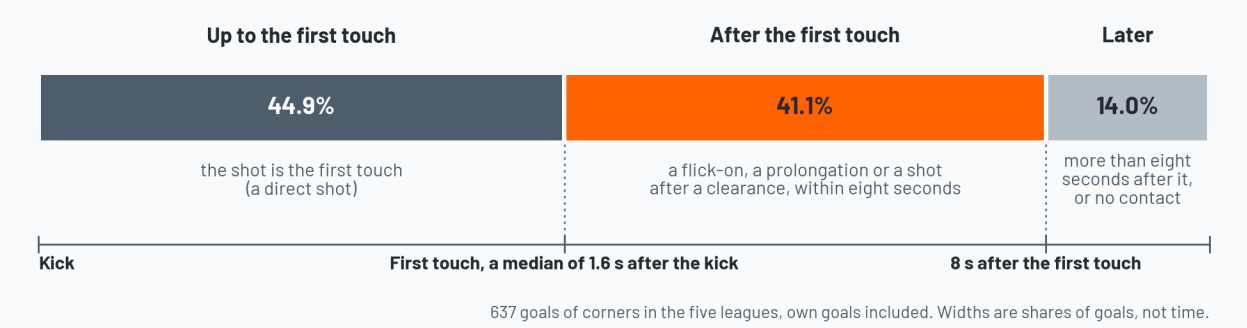


Figure 1: When the goals of a corner arrive, as shares of the goals of corners in the five leagues, own goals included.

Taken together, the results place the value of the second ball in what happens after the first touch rather than in where the players stand at the kick. The contributions are a segmentation of the play after the first touch, built from the event chain alone; a test of whether the freeze-frame at the kick adds anything to the delivery, reported together with the smallest effect the design could have detected; and principles for clubs.

2. Data and method

A corner leaves two traces in the data: the events that follow the kick and the freeze-frame of where the players stood when it was taken. The events can be split at the first touch and counted, which shows how much of what a corner produces comes after it. The freeze-frame can be summarised as the intended plan of each side and set against the delivery in a model, which shows whether that plan decided how the corner ended (Figure 2).

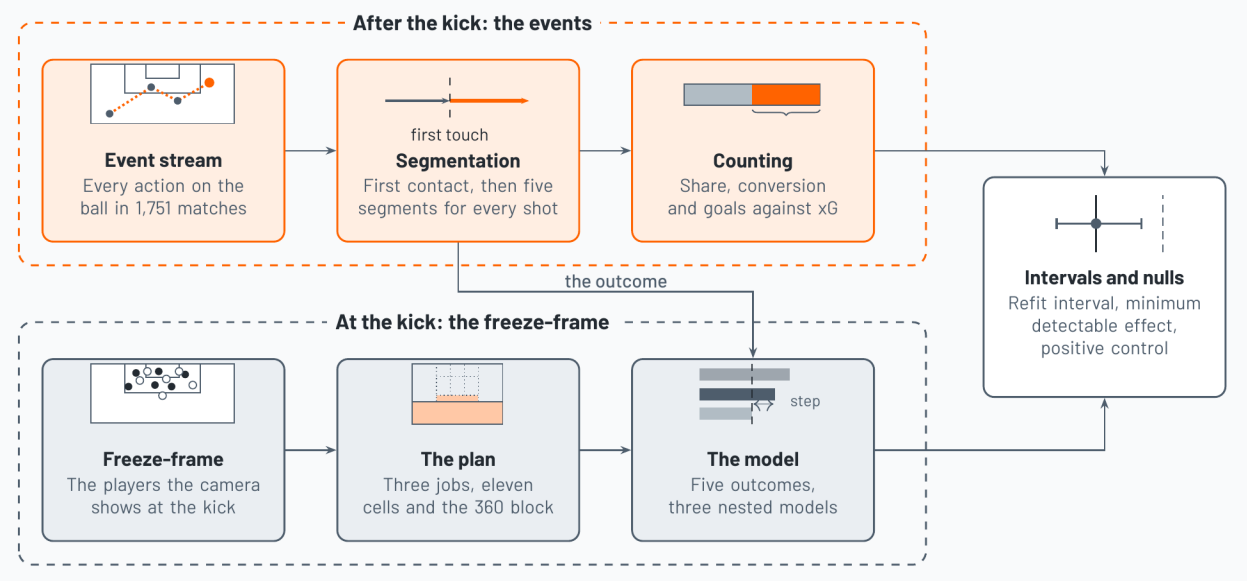


Figure 2: The analysis, from the events and the freeze-frame of a corner to the results.

2.1 Data

The data come from Hudl StatsBomb and cover the 2025/26 season of the five big European leagues, 1,751 matches in all. The event stream records every action on the ball, such as passes, duels, clearances and shots, with its time, location, player, team and outcome, and each shot carries the provider's expected goals (xG) [7]. A 360 freeze-frame comes attached to each event and holds the position of every player the broadcast shows at that instant, tagged as a teammate or an opponent of the player acting. Broadcast tracking at 25 frames per second is available for 30 of the Serie A matches, the home and away games among the six sides that finished top of the league.

On a corner the freeze-frame covers the box and little else, since it only holds the players that the camera shows: on average 18.3% of the pitch is visible, with 16.4 players inside the box, and in the whole season the frames hold a single player in the half of the pitch furthest from the goal being attacked. Whatever needs players far from the ball, such as the rest defence against a counter in the sense of Forcher et al. [8], cannot be measured on these data.

Corners are the main object of study, and free kicks taken from the flank are modelled alongside them. A free kick from the flank reaches the same zones as a corner when it is taken close enough to the goal line, with offside as the only difference. Only those from $x \geq 85$ on the provider's 120-by-80 pitch, outside the width of the box, are kept (Appendix C). Long throws are counted but not modelled, since they leave too few goals in each segment to fit and validate a model.

A corner played short sends the ball to a teammate and not into the box, so the first touch after it is the reception of that pass, which says nothing about a contest. The shots of such a routine would then count as coming after the first touch, with no contest behind it. Its delivery is therefore the first pass that the provider tags as a cross, played by the same side in the same possession within eight seconds of the corner kick (Appendix D). The first contact, the segments and the freeze-frame are taken from that cross, which stands in for the kick. A corner counts as played short when the kick runs along the ground or travels less than fifteen yards, as 3,104 do, and a cross follows in time on 1,326 of them. The rule leaves out where the ball ended, since on a corner that point marks the first contact, so a ground pass driven into the box also counts as played short. When no cross follows in time, the corner has no delivery, and it is counted among the corners taken but left out of the segments and the model, as the long throws are. Table 1 gives the volumes: counts use every delivery, with the corners that have none in a row of their own, whereas the model needs the freeze-frame at the kick and can only use those that carry one (last column).

Table 1: Volumes by population in the five leagues, 2025/26.

Population	Count	Per match	Per side, median	With a 360 at the kick
Corners with a delivery	14,942	8.53	152	86.1%
Flank free kicks, x ≥ 85	2,465	1.41	24.5	92.8%
Long throws, not modelled	1,658	0.95	12	
Short corners, no cross, not modelled	1,778	1.02	16	

2.2 Segmenting the corner at the first touch

The segmentation starts from the first contact, which is the first genuine touch after the delivery by either side: a clearance, a block, a save, a header or an attacker controlling the ball. Rows that record no contact are not touches, so the loser of an aerial duel, a failed reception, a foul, an error flagged by the provider and the keeper's bookkeeping events are left out. Where a first contact is found, the attack wins it on 34.7% of the corners, and it arrives a median of 1.6 seconds after the kick. Figure 3 follows a real corner through the segmentation, Timber's second goal against Leeds United on 23 August 2025, from the kick to the goal. Timber shoots as soon as he receives the ball, which makes that shot direct, and scores later with a shot after a clearance, since Struijk's block had left the ball live.

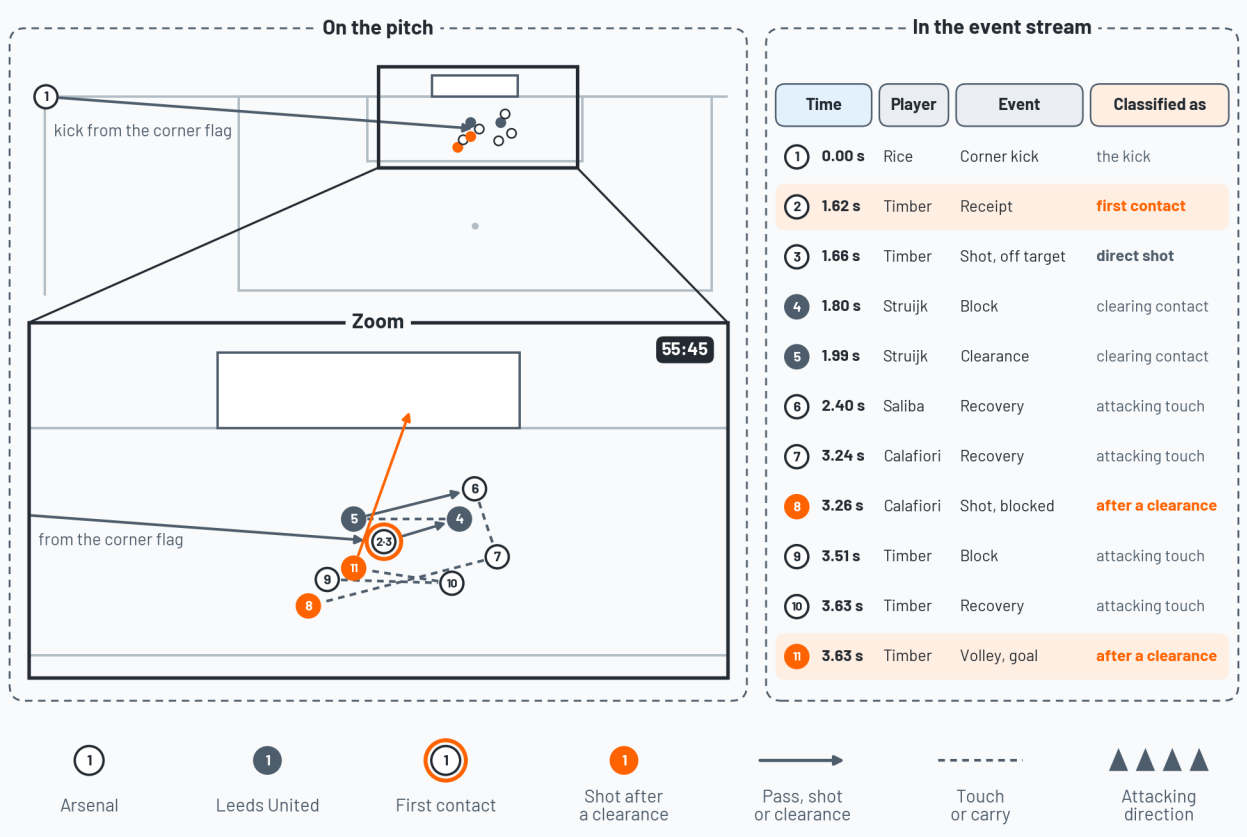


Figure 3: Timber's second goal against Leeds United, 23 August 2025, event by event.

Every attacking shot is then assigned to one of five segments by four yes-or-no questions, answered in the order of Figure 4. A shot more than eight seconds after the first contact, or with no contact at all, is late. A shot that is itself the first touch is direct, and one that follows a defender's touch which left the ball live comes after a clearance. In any other case an attacker kept the ball alive before the shot, and the body part of the last attacking touch decides: a header makes it a flick-on and any other touch a prolongation. Flick-ons, prolongations and shots after a clearance make up what is called *after the first touch* from here on.

The eight seconds are set by the possession (Appendix D). Beyond them, shots begin to come from a different possession than the one the corner started, so a shorter window would drop shots that still grow out of the delivery and a longer one would count open play as part of the corner. Three quantities are computed from the segments: the share of goals after the first touch is the fraction of the goals that corners produce coming from those three segments; the conversion of a segment is its goals per shot; and the rate of a side is its goals after the first touch per corner taken.

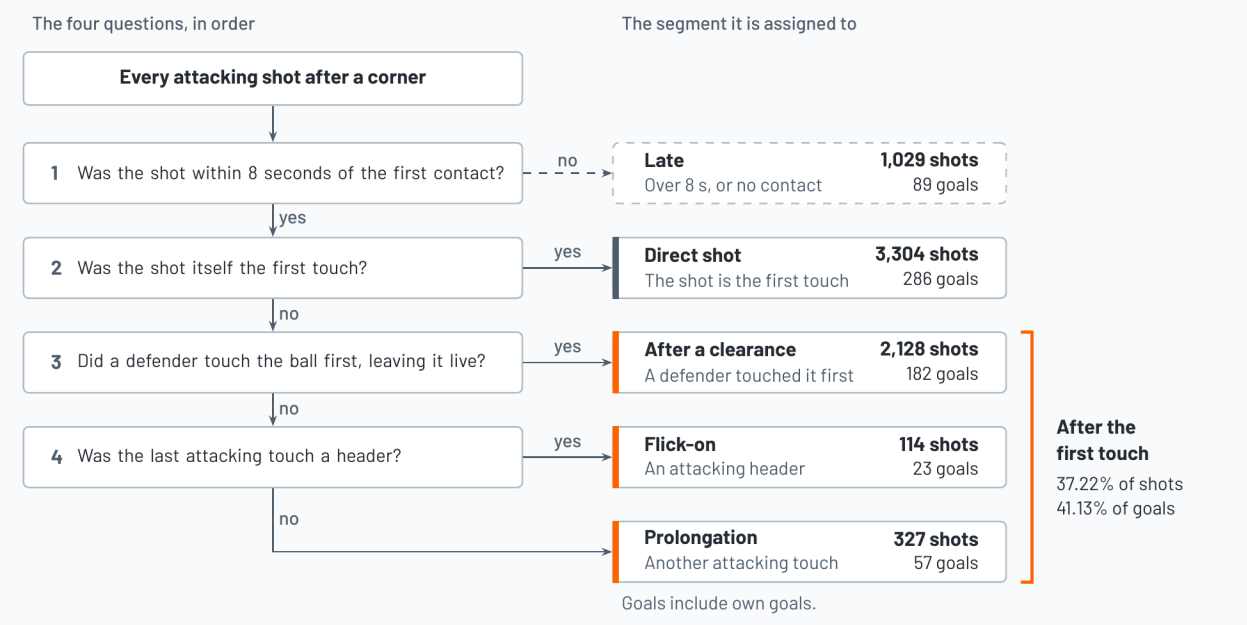


Figure 4: How each corner shot is assigned to a segment, with the shots and goals of each in the five leagues.

The segmentation does not use the provider's phase label or its key-pass annotation, and two checks back this up. Rebuilt from event types, teams and players alone, it gives the same segment as a version built on the key-pass annotation for 94.5% of the shots inside the window, and that annotation shows no sign of depending on whether the shot was a goal (Appendix A). The provider's own phase label agrees with the segmentation on 62.1% of the shots that carry one, and the provider files the direct shot in its first phase 95.0% of the time. Most of the disagreement comes from shots after a clearance that the provider still files in its first phase and from late shots that it keeps in its second phase once the episode has ended. Its label also has no counterpart for the split between flick-ons, prolongations and shots after a clearance.

An own goal counts as a goal in every share of goals. It is not a shot, so it carries no expected goals [9] and stays out of every comparison of goals with xG. It takes the segment of the attacking action that produced it: an own goal off a direct shot counts as direct, and so does a cross that a defender turns in at the first contact.[2] Each shot belongs to the last delivery taken before it, so a corner won from the rebound of another one opens an episode of its own instead of sharing the shots that follow.

2.3 The plan at the kick

The intended plan of each side is inferred from the freeze-frame at the kick, in which each attacker, the taker aside, is assigned one of three jobs from where they stand. Those in the frontal

[2] A penalty won during the episode is not a shot of the delivery, because a foul is not a touch and the kick is taken later as a set piece of its own. Of the 14,942 corners, 69 win one, 47 of them after the first contact. Credited by when the whistle came, the share of goals after the first touch is 41.1% without these penalties and 43.0% with them.

stake the zone where the ball is likely to drop, the frontal being the ground inside the width of the box and up to six yards short of its line, which FIFA's corner study calls the edge of the box [10], together with the central part of the box between its line and fifteen yards from goal. Those elsewhere in the penalty box attack the cross, and the rest hold back, whether they stand deeper or wide of the box. A job is assigned from position alone. The point where the ball ended up plays no part in it, because on a corner that point is where the first contact happened, and a plan built from it would already contain the outcome it is meant to be tested against.

Attackers and defenders are counted in the same eleven cells (Figure 5), the grid of the box that a club reads. Inside the width of the six-yard box, three equal columns cross two rows, the six-yard box itself and the ground from its line to fifteen yards from goal, and the box on either side of that width makes one more cell each. The other three are the edge of the box, which is the frontal, the wide areas and the deep ground further out. Counted one by one, as Figure 5 numbers them, the six-yard box holds three cells, near, central and far, the ground from its line to fifteen yards three more, the two sides of the box two, and the edge, the wide areas of both flanks taken together and the deep ground one each. Every delivery is mirrored, where needed, into one struck from the attackers' left, which makes the near post the same post on every corner. The three jobs are sums of these cells. Together with measures of congestion, marking and the depth of the defensive block, the counts form the 360 block of the model (Appendix A lists its groups; Appendix B gives the origin of the cells).

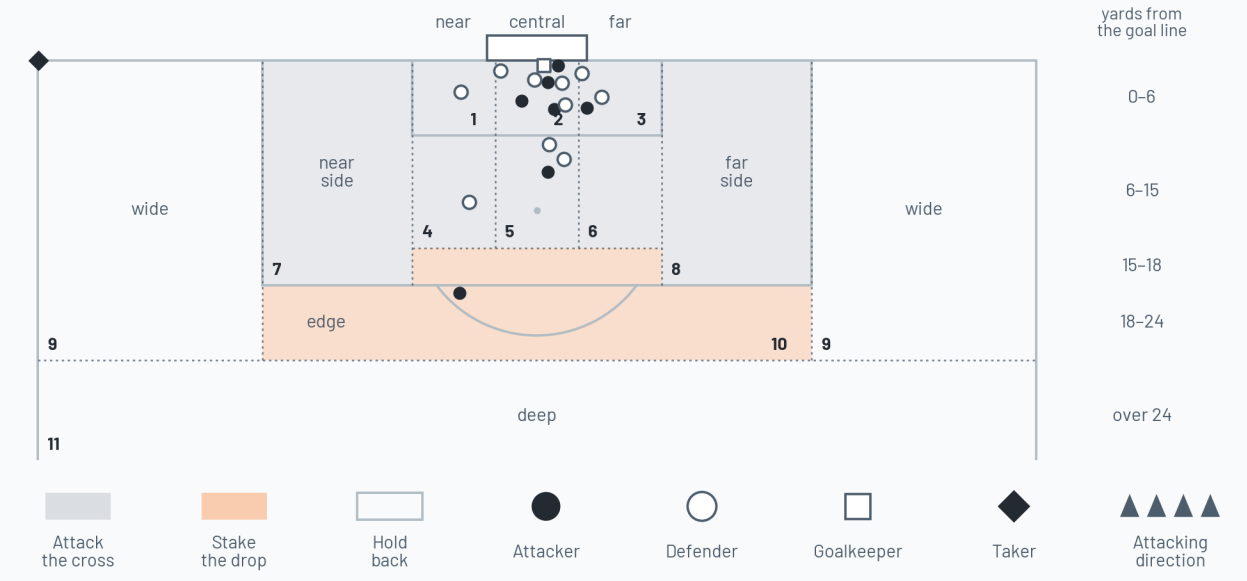


Figure 5: The eleven cells and the three jobs, on the freeze-frame of the corner in Figure 3.

2.4 The model

A corner ends in one of five ways (Table 2): with a shot from the first touch, with a shot after it, with a late shot, with a clearance that no shot follows, or with no second ball at all. The class is set by the first attacking shot, so a corner whose first shot is direct stays in that class even if a rebound is shot later.

Table 2: The five outcomes of a corner and their share of the corners with a freeze-frame at the kick.

Outcome	What happened	Share of corners
Shot from the first touch	The first attacking shot is the first touch	22.5%
Shot after the first touch	The first shot follows a flick-on, a prolongation or a clearance, within eight seconds of the first contact	12.8%
Late shot	The first shot comes more than eight seconds after the first contact, or with no contact at all	5.1%
Clearance, no shot	A defender leaves the ball live and no shot follows	38.3%
No second ball	No shot and no live clearance, so the episode ends without a second ball	21.3%

The five outcomes are modelled with a regularised multinomial logistic regression. For a corner with design *x*, the probability of outcome *k* is

$$p_k(\mathbf{x}) = \frac{\exp\left(\mathbf{w}_k^{\top}\mathbf{x} + b_k\right)}{\sum_{j=1}^{5} \exp\left(\mathbf{w}_j^{\top}\mathbf{x} + b_j\right)} \tag{1}$$

where *x* lists the numbers that describe the corner, the same for every outcome. Each outcome has its own weights *w*, one for each number in *x*, and its own intercept *b*. The weights multiplied by the design and added up, plus the intercept, give the score of the outcome, and the higher the score, the more a corner like this one tends to end that way. The intercept alone is the score of a corner with an average design. The exponential turns each score into a positive number. The sum in the denominator runs over the five outcomes, with *j* from 1 to 5, and rescales the five numbers into probabilities that add up to one. Whatever raises the chance of one outcome therefore takes it from the others.

The weights and the intercepts are fitted by maximum likelihood under an L2 penalty, which pulls the weights towards zero, and the strength of the penalty is tuned in a nested cross-validation. Missing values are filled with the median of the training fold and the columns are standardised on it, so nothing is learned from the matches that are later scored. Folds are grouped by match, which keeps any match from appearing in both training and test, and every probability that enters a comparison is out of fold.

Information is added in three nested models, each containing the one before:
(i) the mix of outcomes alone, which gives every corner the same five probabilities;
(ii) the delivery, which is what is true at the moment of the strike (the side, the swing, the taker's foot, the height of the ball, whether the corner went short or along the ground, and the spot it was

struck from); and
(iii) the delivery plus the freeze-frame at the kick, through the 360 block.

The point where the ball ended is kept out of the delivery on purpose: on a corner it marks the first contact, so including it would leak information about the outcome.

Each model is scored by its log loss, the average surprise of the model at how each corner actually ended. For a single corner it is minus the logarithm of the probability that the model gave to the real ending, so a model that gave that ending a probability of one would score zero, and the less likely it found what happened, the higher its loss. What a model adds to the one before it is the drop in log loss, corner by corner:

$$\ell_i = -\log \hat{p}_i(y_i), \qquad L = \frac{1}{n}\sum_{i=1}^{n} \ell_i, \qquad \Delta = \frac{1}{n}\sum_{i=1}^{n} \left(\ell_i^A - \ell_i^B\right) \tag{2}$$

where the loss of corner *i* comes from the out-of-fold probability that the model gave to the outcome the corner produced, and the step is the mean over corners of the paired difference between a model *A* and the model *B* that adds information to it. A positive step therefore means that the added information helped. The log loss is counted in nats, the unit of natural logarithms, and a step in nats has a simple meaning. A step of 0.01 nats means that the added information raises, on average, the probability that the model gives to what happened by about 1%. Five equal chances would score ln 5 = 1.6094 nats and the mix of outcomes 1.4486, which gives the scale of the steps that follow.

2.5 Intervals and nulls

Intervals are computed by resampling matches, not corners, because two corners from the same match share the teams, the referee and the scoreline, and treating them as independent would make every interval too narrow. Redrawing the matches while keeping each corner's fitted probabilities still leaves out part of the uncertainty, the part that comes from fitting the model on a different set of matches. The whole model is therefore rebuilt inside each of 300 resamples, with the penalty held at its tuned value, and on corners this makes the intervals 1.75 times wider than with the probabilities held fixed.

The refitted draws of a step give three intervals. The percentile form takes the middle 95% of the draws as they fall, the pivotal form reflects them about the observed step, and the normal form keeps the observed step at its centre and takes from the draws only their spread, which is the width the refit adds. The normal form is the one that decides:

$$\begin{aligned} \text{normal}: &\quad \hat{\Delta} \pm \tfrac{1}{2}\,(q_{0.975} - q_{0.025}) \\ \text{percentile}: &\quad [\,q_{0.025},\ q_{0.975}\,] \\ \text{pivotal}: &\quad [\,2\hat{\Delta} - q_{0.975},\ 2\hat{\Delta} - q_{0.025}\,] \end{aligned} \tag{3}$$

where the estimate is the observed step and the q values are percentiles of its refitted draws. Where a claim depends on the form of the interval, the percentile and pivotal intervals are printed beside the normal one (Appendix A).

A difference counts as a finding only if its normal interval leaves zero out. Likewise, a null result is informative only if the method had the power to detect an effect, so it is read against
(i) the smallest effect the design could have detected and
(ii) a positive control that retains the design, the folds and the intervals and changes only the outcome to one where a signal should exist.

For (i), a true step of a given size is planted by recentring the per-corner differences on it, the matches are redrawn, and each redraw's interval is checked for excluding zero. The smallest step that clears zero in eight draws out of ten, multiplied by the width ratio above, is the minimum detectable effect, 0.0087 nats for the 360 block on corners, about 0.9% more probability on what happened.

3. How much is decided after the first touch?

3.1 The share after the first touch

The corner deliveries of the five leagues produced 6,902 shots and 637 goals, own goals included, and the corners played short that no cross reached made 699 more shots and 60 goals, which belong to no segment. What the deliveries yielded came after the first touch in 37.2% of the shots, 40.9% of the expected goals and 41.1% of the goals, however the yield is counted (Table 3). The share rises from the shots to the expected goals because a shot after the first touch is, on average, a better chance than a direct one. Goals are rare, so the share of goals is the least certain of the three, and even its interval lies below an even split between what comes before and after the first touch.

Table 3: The share after the first touch, and its conversion against a direct shot, with 95% intervals over matches.

After the first touch	Corners	Flank free kicks
Share of shots	37.2% [36.1, 38.4]	36.5% [33.3, 39.7]
Share of expected goals	40.9% [38.9, 42.7]	33.8% [29.0, 38.8]
Share of goals	41.1% [37.4, 45.1]	31.5% [21.1, 42.9]
Conversion against a direct shot, pp	+2.0 [+0.4, +3.5]	−0.4 [−4.0, +3.4]

Flank free kicks, a population about a sixth the size, give a different picture. Less of what they yield comes after the first touch, and a shot after the first touch converts no better than a direct one, so the two populations are read apart. A rate pooled over both would be weighted by how many corners and flank free kicks the sample happens to hold.

3.2 Which touches carry the value

A shot after a flick-on scores about two and a half times as often as a direct shot, and one after a prolongation about twice as often (Figure 6). A shot after a clearance does not differ from a direct one (Table 4).

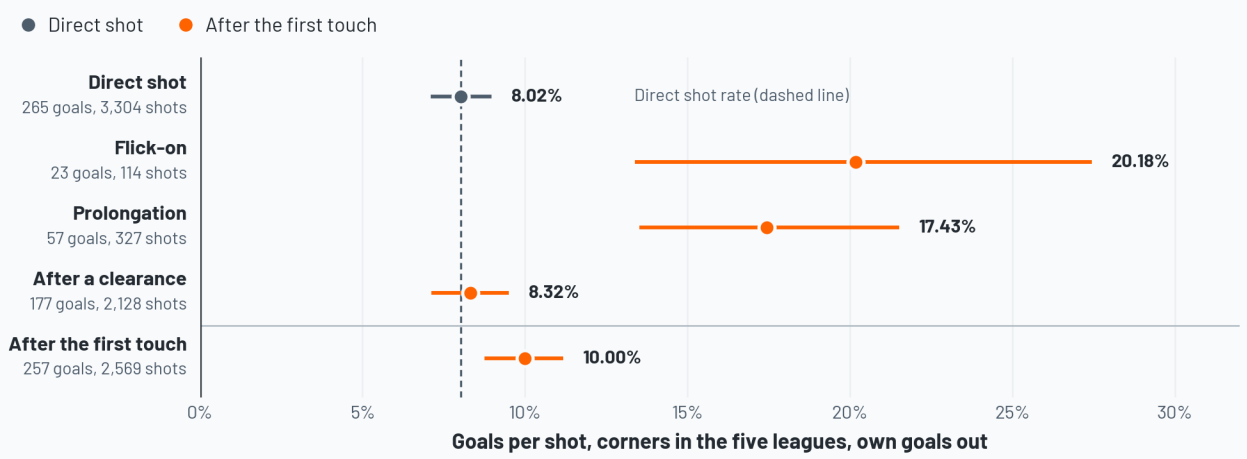


Figure 6: Goals per shot by what came before the shot, corners in the five leagues, own goals out.

Where the last touch before the shot was the attack's, a flick-on or another touch kept the ball alive and the shot came out of the same attack, whereas after a clearance the attack had to win the ball back before it could shoot. Keeping the ball alive is therefore worth more per shot than winning it back, and the extra conversion is what the provider's expected goals predict: after each of the three segments, goals sit close to the goals expected, so these are better chances rather than luckier ones.

Table 4: Corner shots by what came before them, own goals out, with 95% intervals over matches.

Segment	Shots	Goals	Goals per shot	Against a direct shot, pp	Goals over xG
Direct shot	3,304	265	8.02% [7.10, 8.97]		0.87 [0.77, 0.96]
After a flick-on	114	23	20.18% [13.38, 27.45]	+12.15 [+5.36, +19.55]	0.93 [0.65, 1.23]
After a prolongation	327	57	17.43% [13.52, 21.52]	+9.41 [+5.31, +13.56]	0.88 [0.69, 1.07]
After a clearance	2,128	177	8.32% [7.12, 9.51]	+0.30 [−1.26, +1.79]	0.96 [0.84, 1.09]

Two limits come with the flick-on:
(i) it is the thinnest cell, 114 shots and 23 goals, and it cannot be split by league;
(ii) the export records no body part for a duel, so a header won in a duel is filed as a prolongation and the flick-on's shots are a lower bound; since both attacking touches convert above the direct shot, the conclusion does not rest on how the export splits them.

Power et al. [3] found the same ordering on three Premier League seasons of tracking data: corners redirected by a flick-on scored more often than direct ones, and the advantage they reported for winning the second ball in general was not significant.

3.3 Goals against expected goals

The provider's model gives every shot an expected-goals value (xG), and the sum of those values can be set against the goals the shots scored, own goals aside since they carry none. The direct shot scores less often than its xG says, 0.87 goals for each one expected, whereas the shots after the first touch score about what theirs say, 0.94 (Figure 7).

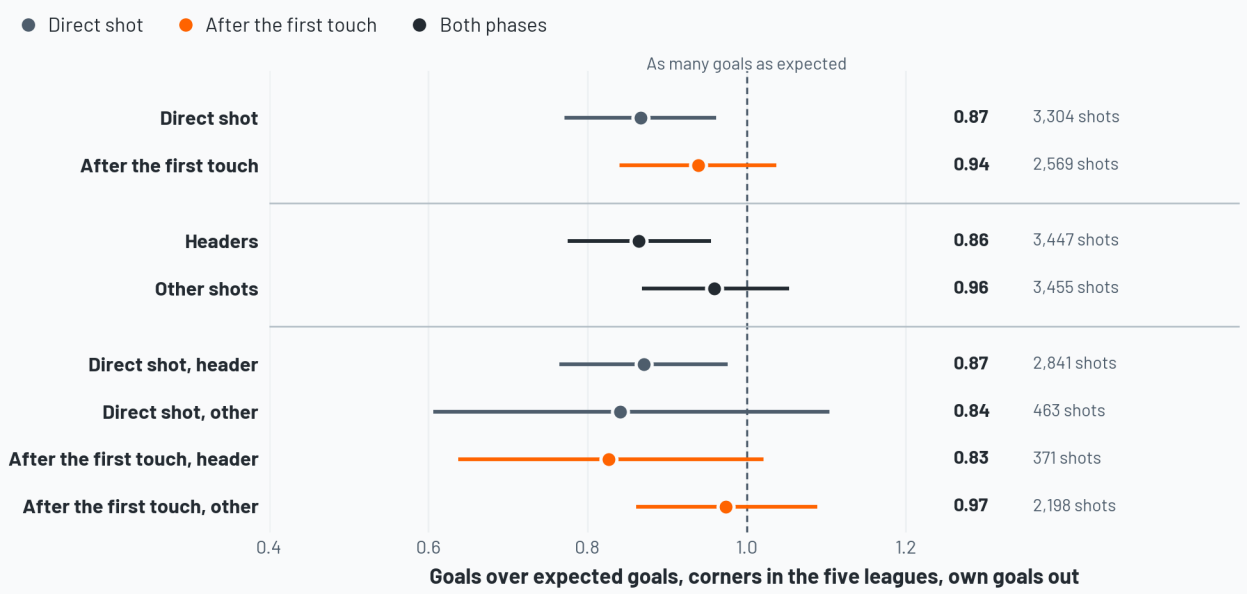


Figure 7: Goals over the provider's expected goals by phase and by body part, corners in the five leagues, own goals out.

The shortfall of the direct shot follows the way it is struck more than the phase, since most direct shots are headers and most shots after the first touch are not. Taken apart, headers score 0.86 of the goals expected of them and all the other shots 0.96, and within headers the phase makes no measurable difference. On corners, then, it is the header that converts below its xG, a feature of how corners convert.

The data cannot tell whether a header from a corner is a worse chance than its features suggest, as the way corners convert would have it, or whether a model that is not built for corners alone places it a little high. Settling that would take a recalibration over open-play headers, and the provider's documentation rules out only that such a shot lies outside the model's range, since the body part and the play context are among its features [7]. Nor does the claim say where in the sequence the header comes. Headers as a group fall below their expected goals under either kind of interval. Those after the first touch, fewer in number, do not under either, whether matches are resampled or an exact Poisson bound is used, which takes the shots as independent.

On flank free kicks the shortfall of the direct shot replicates only narrowly, since its interval just stays below one, and the split by body part shows nothing: neither the headers nor the other shots depart measurably from their expected goals.

4. Was the second ball decided at the kick?

4.1 What the kick adds to the outcome

Over the 12,864 corners that carry a freeze-frame at the kick, knowing how the ball was struck lowers the log loss by 0.0129 nats against the mix of outcomes alone (Table 5), about 1.3% more

probability on the way each corner ended, so the delivery matters. Adding where the players stood at the kick does not improve on it, since the step of the freeze-frame over the delivery is −0.0032 nats and, with the model refitted in each resample, its interval covers zero. Fitted apart and compared corner by corner, a model that knows only the delivery beats one that knows only the 360 block by +0.0089 nats, a lead that clears zero in the normal and the percentile form of the interval, while the pivotal form reaches down to −0.00007. H1 is therefore not supported.

Table 5: What the freeze-frame at the kick adds to the delivery, with 95% intervals over matches.

Measure	Corners	Flank free kicks
Delivery over the mix of outcomes, in nats	+0.0129 [+0.0097, +0.0159]	+0.0003 [−0.0047, +0.0053]
Frame at the kick over the delivery, in nats	−0.0032 [−0.0074, +0.0009]	−0.0174 [−0.0377, +0.0028]
Delivery over the frame at the kick, head to head, in nats	+0.0089 [+0.0038, +0.0141]	+0.0160 [−0.0052, +0.0372]
Groups of features that cover zero once refit	10 of 11	11 of 11
Smallest step detected in eight draws out of ten, in nats	0.0087	0.0339

In Table 5 the first interval holds the fitted probabilities fixed, and the next two rebuild the model in each resample (Section 2.5). The same test can be run group by group, by dropping one group of features at a time. For 10 of the 11 groups, the change in log loss has an interval that covers zero once the model is refitted in each resample (Appendix A lists them). The exception is the delivery, which carries the model. Keeping the plan at the kick therefore buys nothing. Coaching texts hold that attackers who stand close together make tight man-marking almost impossible [11], yet dropping the congestion group costs the model only 0.0001 nats, with an interval that covers zero.

Dropping a group changes at once the information, the width of the design and the regularisation, so a second control shuffles the group across corners instead, which removes only the information it carries [12]. A shuffle costs something even when a group carries nothing, so each group is read against 30 blocks of pure noise as wide as itself, whose 95th percentile at the width of the delivery is 0.0008 nats. Shuffling the delivery costs 0.0075 nats, about ten times that. Of the 360 groups only marking and congestion cost more than every block of their width, 0.0012 and 0.0012 nats, and dropping them costs 0.0004 and 0.0001 nats, each with an interval that covers zero, so neither shuffle counts as information.

Of the corners in the model, 9.8% are the cross of a corner played short. Scored at a single regularisation, which leaves only the rows to change, the lead of the delivery over the 360 block is 0.0088 nats over every corner of the model and 0.0080 over the long ones alone, so the crosses do not carry it. Counting each short corner at its own kick, where the first touch is the reception of its own pass, the lead would be 0.0312 nats, 3.6 times as large. The crosses alone beat the 360 block

when they come within eight seconds of the corner kick, by 0.0649 nats, and at the three other windows measured, five, ten and fifteen seconds.

On flank free kicks, by contrast, the delivery design does not beat the mix of outcomes (+0.0003 nats), every one of the 11 groups of features covers zero once refit, and the design could not have detected a step smaller than 0.0339 nats, so a null on that population describes the sample.

4.2 What the design could have detected

On corners, the most that the freeze-frame at the kick could be adding, the upper bound of its step, is +0.0009 nats, far below the 0.0087 nats that the design detects in eight draws out of ten. An effect of a size that the design can detect was therefore not there. The bound is also small beside what the delivery adds. Taking the lower bound of the delivery's own step, 0.0097 nats, as the reference and a tenth of it, 0.0010 nats, as the margin [13], the frame's upper bound comes to 9.2% of the reference and stays inside the margin. That verdict depends on the form of the interval, since in pivotal form the bound is +0.0066 nats and the margin is exceeded. The margin did not move once the bound was known, but I chose the normal form after the pivotal verdict, so both are given.

On who wins the first contact, an outcome that the bodies at the kick should predict, the delivery saves 0.0170 nats over the base rate, so this positive control shows that the method detects an effect when there is one. On the long corners alone the saving is 0.0109 nats, smaller and clear of zero, and with the outcome shuffled across corners the same design gains nothing. The freeze-frame at the kick adds nothing to the delivery on this target either, −0.0003 nats over every corner and −0.0008 over the long ones, so the order found for the outcome of the corner comes back on a target where the method detects signal. On flank free kicks the control finds none, the delivery gaining +0.0001 nats over the base rate and the whole 360 block −0.0049, both within noise, so a null on that population cannot be set against the control as the corner nulls are.

A null on average could still hide a gain on some corners, so a single test in the manner of Chernozhukov et al. [14] was made before any subgroup was opened. It finds that the difference between the two designs is predictable, and a placebo that permutes the inputs of the prediction finds nothing. But the gap between the fifth of corners predicted to gain most and the fifth predicted to gain least is 0.0089 nats and does not differ from zero ($p = 0.11$). The difference varies without saying for whom, so no subgroup is opened. On flank free kicks no structure is detected at all.

4.3 What the frame at the kick does carry

On the corners that the attack reached first, where the point at which the ball ended is where it was aimed, the attackers' positions at the kick predict the zone of the box that the ball went to, beyond how it was struck: the log loss of the zone falls by 0.0234 nats (Table 6). The idea was developed on the Premier League, which is therefore kept apart, and the claim is made on the other four leagues pooled, 9,247 long corners in all. No league resolves it alone in the direction of the pooled sample, and Ligue 1 clears zero in the other, so it is never made league by league. Putting the league into

both designs, to rule out that the frame is only naming the league, leaves a saving of 0.0186 nats that is still clear of zero on the fixed fit.

Table 6: The aim read from the frame at the kick, in nats saved, and the near-post discount of two maps, in xG per corner, with 95% refit intervals.

Comparison	Estimate	Normal	Pivotal	Percentile
Aim: saved by the frame at the kick	+0.0234	[+0.0042, +0.0425]	[+0.0161, +0.0544]	[−0.0077, +0.0307]
Near-post discount, end-zone map	−0.0140	[−0.0233, −0.0047]	[−0.0225, −0.0039]	[−0.0241, −0.0055]
Near-post discount, aim map	−0.0015	[−0.0108, +0.0078]	[−0.0085, +0.0101]	[−0.0130, +0.0055]
Difference between the two maps	−0.0125	[−0.0255, +0.0005]	[−0.0261, −0.0001]	[−0.0249, +0.0011]

The interval of the aim excludes zero in normal and in pivotal form and covers it as percentiles, because resampling puts a shift into every draw (Appendix A). It matters for the map that a club would draw to choose a target, which values corners by the zone where they end. On that end-zone map, corners that end up to the near post in the six yards in front of goal are worth 0.0140 xG per corner less than those that end beyond the far post. But where a corner ends is where the first contact happened, and when the defence made it that is where a defender met the ball: the end of a failed pass is where it was intercepted, not where it was meant to go [15]. Split by who won the first contact, the contrast is +0.0113 where the attack made it and +0.0014 where the defence did, neither distinguishable from zero, so the discount is absent on both sides. It appears only when the two are pooled, because the near-post end zone fills with corners that the defence reached first, which are worth less. The difference between the two maps clears zero in pivotal form only, and narrowly, since in normal form it reaches +0.00051 (Table 6). The end-zone map therefore values the interception. The aim map, which groups the corners by the aim inferred from the frame at the kick, shows a near-post discount that cannot be told from zero, and neither map says where to aim.

Hence the plan at the kick carries intent, and intent adds nothing to how the corner ends once the delivery is known. On flank free kicks neither the aim nor the bias of the map is resolved, over 1,830 deliveries.

4.4 Who wins the second ball

The second ball is won by the side that makes the next touch after the clearance, within the eight seconds of the episode. On the 7,558 corners with a freeze-frame at the kick whose second ball the event chain resolves, the attack wins it 70.0% of the time, so always naming the attack would be right seven times in ten. The delivery adds nothing measurable to that base rate, +0.0006 nats saved. Where the players stand at the kick does a little better, and not enough (Table 7). Its

geometry saves 0.0024 nats over the delivery, with an interval that covers zero in the normal and the percentile form and clears it in the pivotal form, so the verdict is borderline, and in any of the three forms the gain is too small for a coach to act on. A gradient-boosted model fitted on the same columns does no better, and neither does a race in which every player runs from their position at the kick to the median point where cleared balls land.

The first contact comes a median of 1.6 seconds after the kick, and on 6,193 of the 7,558 corners it is the clearing touch itself. A frame at that moment exists for 7,413 of the 7,558, and on those corners the same geometry, measured at the first contact instead of at the kick, saves 0.0065 nats over the delivery, about 0.7% more probability on who wins the second ball. That interval leaves zero out in all three forms, and also after correcting for the comparisons made at the first contact, so at the first contact the positions do add to the delivery. The rule fixed for these comparisons requires that the pivotal interval leave zero out and that its decisive bound keep its side in each of five repetitions of the resampling, and this one meets it. The geometry at the first contact also beats the same geometry at the kick, by 0.0040 nats, and its interval leaves zero out in all three forms, by +0.0002 at the nearest bound, but in one of the five repetitions its pivotal bound crosses zero, so the rule leaves that comparison undecided, and after the correction for the comparisons its interval covers zero as well. Where the bodies stand adds to knowing who touched the ball, when and where, by 0.0054, in all three forms and after the correction, and its decisive bound keeps its side in every repetition. The three comparisons keep their direction when the attackers beside the corner spot, among whom the taker cannot be told apart in a 360 frame, are removed. The same race, run from the positions at the first contact, does not clear zero.

The first of these is the clearest positive result of the study, and it is small, below the 0.0067 nats that the design detects in eight draws out of ten. It also arrives too late to be planned. The estimate is conditional on information that arrives after the decision a coach makes, so what it values is a camera's view at the first contact and not a routine, and nothing from that frame may enter a model of the kick. On flank free kicks the frame at the first contact beats the delivery in none of the comparisons, over 1,169 free kicks.

Figure 8 puts the three outcomes side by side. The delivery carries the outcome of the corner and the first contact, the frame at the kick adds to neither, and on the second ball it is the frame at the first contact that clears zero, below the line that marks the smallest effect the design detects in eight draws out of ten.

Table 7: Who wins the second ball, read at the kick and at the first contact, as log loss saved in nats, with 95% refit intervals in three forms.

Comparison	Saved	Normal	Pivotal	Percentile
Frame at the kick, over the delivery	+0.0024	[−0.0006, +0.0053]	[+0.0011, +0.0070]	[−0.0022, +0.0036]
Frame at the first contact, over the delivery	+0.0065	[+0.0030, +0.0101]	[+0.0044, +0.0116]	[+0.0015, +0.0087]
Frame at the first contact, over the frame at the kick	+0.0040	[+0.0002, +0.0077]	[+0.0002, +0.0077]	[+0.0003, +0.0077]
Bodies at the first contact, over who touched the ball and where	+0.0054	[+0.0020, +0.0088]	[+0.0037, +0.0104]	[+0.0003, +0.0071]

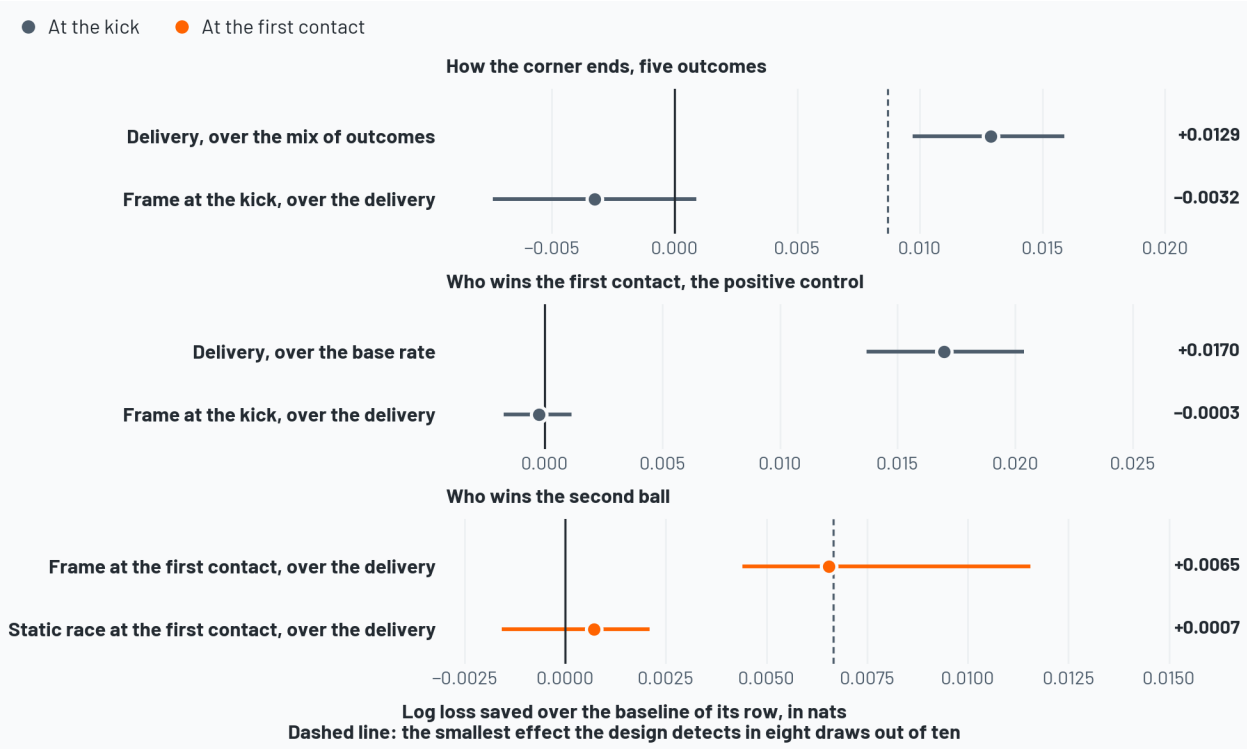


Figure 8: What each piece of information adds, as log loss saved over the baseline of its row, for three outcomes of a corner.

5. What can a club do about the second ball?

5.1 Plan or race

A set-piece analyst's account of Arsenal's corners puts two rebound players on the edge of the box, with a security player behind them [16], which is the job that Section 2.3 calls staking the drop. Once a defender has cleared the ball, someone makes the next touch, and the plan either had a player in the frontal within ten yards of where that touch met the ball or it did not. Crossing the two facts gives four states (Figure 9). The second ball is *won as planned* when the attack made the touch with a player staked near the drop and *rescued* when it made it with nobody staked; it is *staked but lost* when the defence made the touch despite a staked player and *conceded open* when it made it with nobody there. The states need both a freeze-frame at the kick and a player who contests the loose ball, so they are known on 7,558 of the 9,175 cleared corners, and most of the second balls that the attack wins are rescues, 4,084 of 5,288.

H3 expected a planned win to be worth more than a rescue, and once the drop is held equal it is not (Table 8). 21.0% of second balls are followed by a shot, worth 0.0798 xG on average. The xG that follows a second ball won with a player staked near the drop is 0.0070 more than after one won without, and the interval of that difference leaves zero out. That comparison, however, sets the two states against each other wherever their drops happen to fall, so the premium is also tested at the same position, within cells of the pitch of four sizes and with a regression on the drop point, and there it is gone, since neither a cell size nor the regression brings it back. A second ball won as planned is worth more because of where it drops, and at the same drop a player staked there adds nothing measurable to what winning it is worth. When the attack loses the ball instead, the xG that follows a staked drop, 0.0037, is lower than after an unstaked one, 0.0073, and on flank free kicks the gap points the same way without clearing zero. Because staking is a choice that teams make, these are differences between states rather than causal effects.

Table 8: What the plan is worth after a cleared corner, in xG per second ball, with 95% intervals over matches.

Comparison	Corners	Flank free kicks
Won as planned minus rescued	+0.0070 [+0.0026, +0.0116]	+0.0152 [+0.0050, +0.0264]
Won as planned minus rescued, within cells of six yards	+0.0016 [−0.0028, +0.0066]	−0.0080 [−0.0226, +0.0078]
Won as planned minus rescued, by regression on the drop point	−0.0029 [−0.0077, +0.0021]	−0.0040 [−0.0177, +0.0096]
Staked but lost minus conceded open	−0.0036 [−0.0063, −0.0009]	−0.0045 [−0.0129, +0.0023]

Whether someone reacted faster on the day can be followed on the tracked matches. The events and the tracking keep separate clocks, and in every one of the 30 matches the tracking runs behind, by a median of 1.13 seconds, so each half is first aligned on where the events place the ball and where the tracking detects it. On the aligned clock the ball sits a median of 2.0 m from the corner flag when the corner is taken. For each of the 117 cleared corners whose second ball the

event chain resolves, the drop is where the next touch met the ball, and the side that won the second ball either already had the player nearest that point at the kick, had one get nearest while the ball was in the air before the clearance, or won the run that followed the clearance.

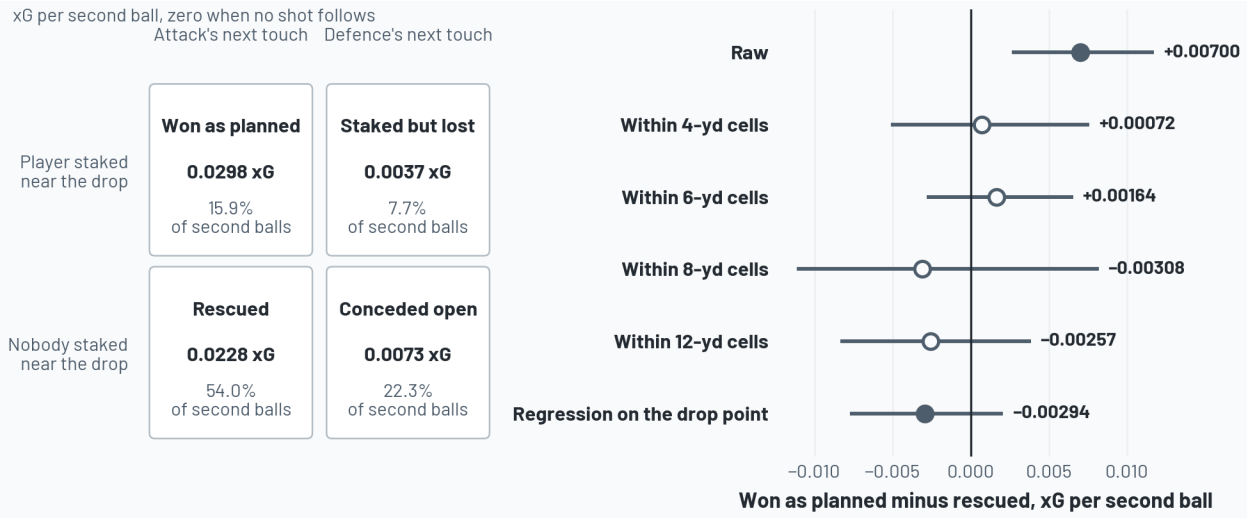


Figure 9: The four states of a cleared corner, the xG per second ball (zero where no shot follows) and the plan's premium.

The side already nearest at the kick won 77.8% of these second balls (Table 9), 21.3 pp more than if being nearest at the kick had nothing to do with winning. A side that got nearest during the flight won 6.8%, and the run after the clearance decided 15.4%. When the attack won the second ball it was already nearest at the kick 88.2% of the time, whereas when the defence won it the run decided 31.7%. Among the players near the drop whom the tracking detects throughout, the defenders started moving towards it 0.14 seconds before the attackers on average, a difference that covers zero, so neither side reacted sooner (Figure 10).

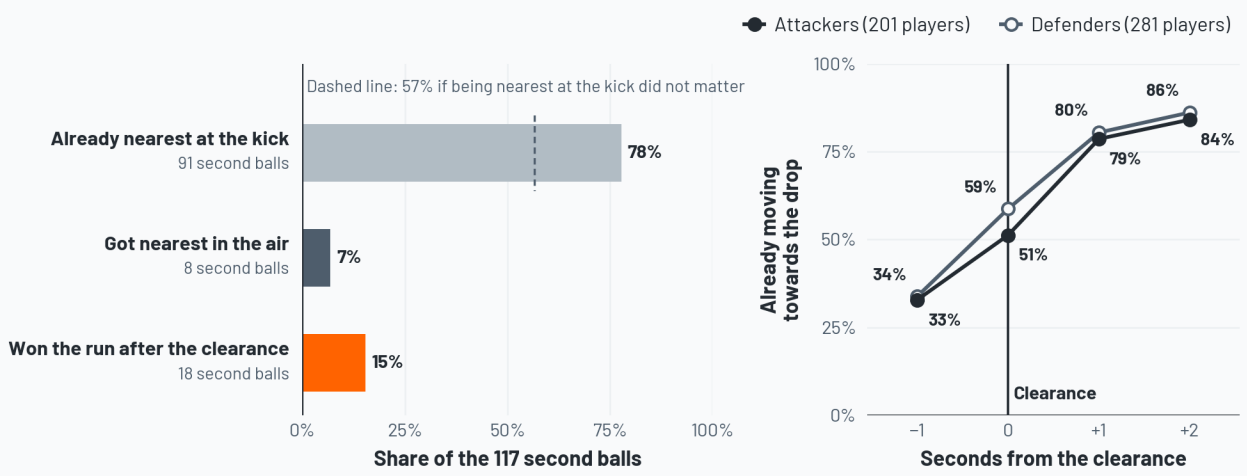


Figure 10: How second balls were won at tracked corners (left) and the share of players near the drop already moving towards it (right).

The three cases are counted after the fact, with the drop known, so they describe how these second balls were won and predict nothing. Beside Section 4.4 they place the uncertainty of the second ball in where it drops, which the kick does not tell, rather than in who reacts to it. They come from thirty matches between six top sides, and the onsets rest on the 48% of the players near the drop that the tracking detects throughout.

H2 asked whether a model of who can reach the loose ball first [17] beats the static picture of where the bodies stand. As a forecast it cannot be tested on thirty matches, and the cases above do not test it, although they lean against it, since the run after the clearance decides few of these second balls.

Table 9: How the second balls of the tracked corners were won, with 95% intervals over matches.

Measure	Estimate	95% interval
Won by the side already nearest the drop at the kick, %	77.8	[70.2, 85.5]
Won by the side that got nearest during the flight, %	6.8	[3.0, 11.1]
Won by the side that won the run after the clearance, %	15.4	[9.4, 21.1]
Already nearest at the kick if that had nothing to do with winning, %	56.5	[52.4, 63.2]
Onset towards the drop, defenders minus attackers, in seconds	−0.14	[−0.32, +0.01]

FIFA's study of the 2025 Club World Cup found a second phase in 57% of the corners defended with no player designated to the edge of the box and in 40% of those defended with one [10]. The same question can be put to these data, with a defender in the frontal at the kick as the treatment, but the corners that had one, 40.2% of them, differ so much from the rest that 57.8% of the propensity scores fall outside 0.05 to 0.95, and none of the nine specifications tried leaves zero out. The effect

is not estimable, whatever interval an estimator returns. FIFA's figure is also a share of the corners that reached a second phase, whereas the class modelled here is a share of all deliveries, so the two cannot be set side by side. The counter that the attack concedes cannot be answered either, because almost no defending shot inside the window stays in the possession that the clearance started: at eight seconds only 5 defending shots follow a clearance, and 40.0% of them are in that possession. Past twelve seconds the share falls with every step, as the window begins to count the other team's next attack.

5.2 Teams

Arsenal's rate of goals after the first touch is 0.037 per corner, and Tottenham's 0.057. A model of teams within leagues discounts the sampling noise, which shrinks both rates towards their leagues (Table 10), and Arsenal's own effect is then no longer distinguishable from its league's. What survives is a probability of 0.31 of being among the top five of the 96 sides, with a rank anywhere from 1 to 56: better than most, and nowhere near a ranking. Tottenham's case is of the same kind (Table 10), and the probability that Arsenal's rate is above Tottenham's is 0.23, which leans towards the side that finished 17th without telling it apart from the champion.

Table 10: Arsenal and Tottenham after the first touch, as counted and with the sampling noise discounted, with 95% intervals.

Measure	Arsenal	Tottenham Hotspur
Corner goals after the first touch	8	12
Rate per corner, as counted	0.037	0.057
Rate per corner, noise discounted	0.024 [0.013, 0.039]	0.032 [0.015, 0.053]
Probability of being among the top five	0.31	0.70
Median rank	10 [1, 56]	3 [1, 31]

More generally, within a league the spread between teams stays inside the 95% band that binomial noise would give, though close to its upper bound (one-sided p = 0.039, against the 0.025 that the bound marks). The reliability of the rate between two halves of the season, projected to a full season, is 0.29, poor on the usual bands [18] and with the whole of its interval below 0.50, so a single season cannot rank teams on this rate. The number of corners a side wins per match, the control, reaches 0.74, so the method does find a property of teams where there is one.

The top two of each league win more corners, 6.14 per match against 4.62 for the rest (Figure 11), yet what each corner gives them after the first touch does not follow. Their expected goals per corner after the first touch, 0.0173 against 0.0162, are not resolved: the interval of the gap covers zero and reaches past a quarter of the pooled rate, the margin that I set for calling two groups alike, so the data neither show a gap nor rule one out. On this measure there is no sign that the channel after the first touch belongs to the top sides, though a gap of up to about a third of the rest's rate is not excluded.

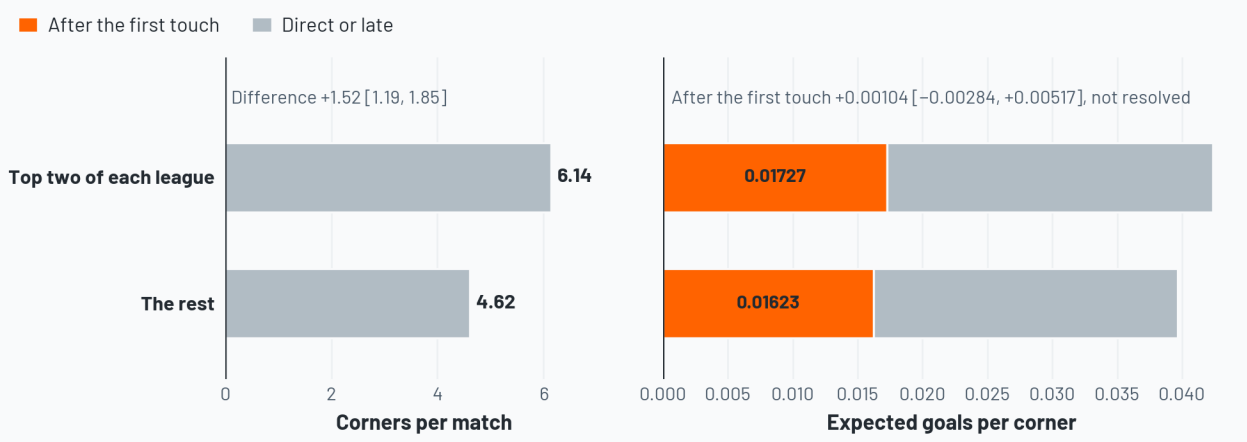


Figure 11: Corners per match (left) and expected goals per corner (right) for the top two sides of each league and for the rest.

5.3 Principles for clubs

Six principles follow from the results. The data do not test these.

1. Count the second ball in what a corner is worth. A routine judged on the first contact alone leaves 41.1% of the goals that corners produce unexplained.

2. Keep the ball alive with an attacking touch. A shot after a flick-on converts 12.2 pp more often than a direct shot and one after a prolongation 9.4 pp more, while a ball won back after a clearance scores like a direct shot.

3. Expect less from the direct shot than its chance quality suggests. It converts at 0.87 of its expected goals, and the shortfall follows the header more than the phase.

4. Explore second-ball plays instead of taking the current set-up as settled. Once the delivery is known, where the attackers stand at the kick adds nothing measurable to how the corner ends, and at the same drop a player staked there does not make winning the second ball worth more, but one season, unevenly covered across leagues and seen through a freeze-frame that holds little more than the box, cannot show that nothing can be planned.

5. Watch the second ball at the first contact, not at the kick. The frame taken at the first contact, a median of 1.6 seconds after the kick, adds to the delivery in telling who wins the second ball, though the gain is small, and the data do not decide whether it adds more than the frame at the kick. It arrives after the decision a coach makes, so it cannot be rehearsed.

6. Judge a side's second balls over several seasons. One season cannot rank teams on them.

6. Related work

Power et al. [3] followed set pieces beyond the first contact. Over three Premier League seasons of tracking they segmented each set piece with what they call a "grammar" and tested ten common

beliefs one by one, among them that a corner is more likely to be scored on the second ball than directly. They marked that belief confirmed although its p-value was 0.86, a non-difference, and Section 3.2 finds the same non-difference for the shot after a clearance while it replicates their flick-on result. Mitrotasios et al. [19] coded a "second play" (the attack kept or restarted after the corner) on 3,620 LaLiga corners by observation, and it was associated with the corner ending in an attempt but not with it ending in a goal. Litwitz et al. [20] coded the rebound goals of seven Premier League seasons, which they found tied to a crowded box and more common after corners. Their rebound is only one of the routes that this paper files as a shot after a clearance, and their unit is the goal, so none of their shares is a rate per corner. The provider's own set-piece model labels a second phase in the event data [21], so the label is not new either, and Lamberts proposes an event-based measure of second-ball retention in open play, about keeping the ball rather than the chance it creates [22]. What none of these works asks is what the positions at the kick say about the second ball.

Closer to that question, Shaw and Gopaladesikan [4] classify each defender as a man or zonal marker from tracking data and describe each attacking run by where it starts and where it ends, at most a second after the first touch. TacticAI [5] predicts, from the frame in which the corner is taken, who receives the first contact and whether a shot follows; it also suggests adjusted setups, which experts preferred. Groom et al. [6] infer marking roles without labels and report that club analysts, shown a snapshot of the delivery, had the player of the first contact among their top three picks only 23.75% of the time. In a second study they train a graph network to move the attackers so as to raise the probability of a shot from the first contact against a defence held fixed [23]. None of these follows the ball after it drops.

Models that value what follows an action are older, and VAEP values the actions on the ball by how they change the chances of scoring and of conceding, while stating that defending, which is more about positioning and anticipation, falls outside what it values [24]. The Bundesliga's Set Piece Threat, the deployed baseline for set pieces, scores the shots that a corner or a free kick resulted in with fixed weights on their outcome and their xG, with no window and no positions of its own [25]. Causal studies of tactical choices are rarer: the one at scale applies double machine learning to formations over more than 22,000 matches and finds no major isolated impact of formation [26].

In the 2025 edition of the Hudl Performance Insights competition, Cascioli et al. [27] valued off-ball defending from StatsBomb 360 freeze-frames and left set pieces out, and Greenhalgh et al. [28] modelled the defence of the long throw with an interpretable model on the frame of the throw and a sequence model of the events after it, second-ball contests among them. The contribution here is narrower than either: a value for the second ball at corners, as competing outcomes modelled from the freeze-frame at the kick, set against a baseline that knows only how the ball was struck and checked by a control that shows what the method can detect.

7. Limitations and future work

Telling the share of Arsenal's corner goals that came after the first touch apart from the share of the five leagues, with 80% power, would take 1,753 corner goals, and 33 for Tottenham. The two scored 19 and 18 in the season that broke the record, so a single season is the main limit. The freeze-frame at the kick shows the box and not much more, and it exists for only some of the corners, 73.7% of the Premier League's against 92.2% of Serie A's, so the corners that enter the model come more from some leagues than from others. Of the corners played short, the model holds only those that a cross reaches within eight seconds of the corner kick, so what it says of them describes the routines that end in a cross. A corner driven along the ground into the box counts as played short and has no delivery unless a cross follows. A later ball into the box that the provider does not tag as a cross is not a delivery either. The tracked second balls come from thirty matches between six top sides, and a broadcast feed shows only part of the play [29], so the onsets of Section 5.1 rest on the players it detects throughout. Flank free kicks are thinner still, and there the positive control does not detect signal either, so their nulls describe this sample rather than the game. The states of the plan are descriptive, since staking is a choice, and the effect of a defender at the edge of the box and the counter that a clearance concedes cannot be estimated (Section 5.1). The cells of the box are the grid that Opta Analyst draws over the box, the one clubs use (Appendix B gives their origin). Last, the intervals refit the model inside every resample but were not checked against nested cross-validation [30], the standard that would measure their width directly. Since each draw fits on about two thirds of the matches, they are expected to err wide.

Four directions follow:

1. More tracking, over more leagues and sides, would show whether the second ball goes to the side already nearest the drop beyond six top sides, and would let it be read as a forecast, with the drop predicted at the kick rather than known.

2. Several seasons of one league would show whether a side's rate after the first touch is a property of the side, and by the Spearman-Brown formula four seasons of this volume would lift the reliability of the rate from 0.29 to about 0.62, and six to about 0.71.

3. One hypothesis is worth following, that Arsenal and Tottenham may be early. If clubs are only starting to work on the second ball, later seasons should show the rate becoming a property of the sides that train it, and patterns in the plays may appear as they do for the aim of the first ball (Section 4.3). Estimating what another set-up would have yielded needs the decision read at the first contact, more leagues, and an off-policy evaluation with a positive control, which is the step from these measurements to a tactical theory of second-ball plays.

4. A side that lets the second phase run loose may surprise opponents who have prepared for the delivery, and the uncertainty may cost them more than it costs the side. Cascioli et al. [31], who value giving the ball away on purpose near the opponent's goal, end with a mixed policy that picks the option from the game context, "possibly with some randomness for game-theoretic reasons". These data cannot test it, since they carry no measure of how well an opponent has scouted a

routine, and the congestion result of Section 4.1 gives no sign that a side can produce the loose ball on demand.

8. Conclusion

A corner is prepared up to its first touch, yet 41.1% of the goals that corners produce come after it, and per shot the value sits in the attacking touch that keeps the ball alive. At the kick, what the data show to matter is how the ball is struck. Once the delivery is known, the positions of the players add nothing measurable to how a corner ends, and, with the drop held equal, a player staked where the ball will drop makes winning it worth no more, which describes the second balls of one season and not what a side that prepares them could do. What adds to the delivery in telling who wins it is the frame taken at the first contact, when the plan has played out and can no longer be changed. Who reacted faster can be counted only on thirty tracked matches, where the side already nearest the drop won most second balls and neither side reacted faster. For a club the consequences are modest (Section 5.3): count the second ball in what a corner is worth, keep the ball alive with an attacking touch, and do not rank sides on one season of second balls. For analysts the next step is a season of tracking that sees the ball when it drops. For coaches, two questions stay open:

1. Is the second phase, which no one can script, a risk to reduce or a source of uncertainty to use?
2. Would a tactical theory of second-ball plays make it something a coach can prepare for?

References

[1] Blitz, S. (2025, August 23). Arsenal's goals from corners: How can Premier League rivals stop Mikel Arteta and Nicolas Jover's set-piece tactics? *Sky Sports*. https://www.skysports.com/football/news/11670/13415710/arsenals-goals-from-corners-how-can-premier-league-rivals-stop-mikel-arteta-and-nicolas-jovers-set-piece-tactics
[2] Tweedale, A. (2026, May 1). In a Premier League season dominated by set-pieces, corners could decide the title race and relegation battle. *The Analyst*. https://theanalyst.com/articles/arsenal-tottenham-corners-record-premier-league-title-race-relegation-battle
[3] Power, P., Hobbs, J., Ruiz, H., Wei, X., & Lucey, P. (2018). *Mythbusting set-pieces in soccer*. MIT Sloan Sports Analytics Conference, Boston, MA, United States.
[4] Shaw, L., & Gopaladesikan, S. (2021). *Routine inspection: A playbook for corner kicks*. 15th MIT Sloan Sports Analytics Conference, virtual.
[5] Wang, Z., et al. (2024). TacticAI: An AI assistant for football tactics. *Nature Communications, 15*, Article 1906. https://doi.org/10.1038/s41467-024-45965-x
[6] Groom, S., Belo, F., Rice, A., Anderson, L., & Wang, S. (2026). A machine learning framework for off-ball defensive evaluation applied to corner kicks. *npj Artificial Intelligence*. https://doi.org/10.1038/s44387-026-00138-7
[7] Hudl StatsBomb. (2022). *Expected goals 2.0* [White paper].
[8] Forcher, L., Forcher, L., Altmann, S., Jekauc, D., & Kempe, M. (2023). The success factors of rest defense in soccer: A mixed-methods approach of expert interviews, tracking data, and machine learning. *Journal of Sports Science and Medicine, 22*, 707–725. https://doi.org/10.52082/jssm.2023.707
[9] Anzer, G., & Bauer, P. (2021). A goal scoring probability model for shots based on synchronized positional and event data in football (soccer). *Frontiers in Sports and Active Living, 3*, Article 624475. https://doi.org/10.3389/fspor.2021.624475
[10] FIFA Football Performance Insights Team. (2025, September 24). *Set plays: Defending corner strategies* [FIFA Club World Cup 2025 technical analysis]. FIFA Training Centre. https://www.fifatrainingcentre.com/en/game/tournaments/fcwc/2025/team-analyses/set-plays-defending-corner-strategies.php
[11] IB. (2019, December 6). Tactical theory: Set-pieces. *Spielverlagerung*. https://spielverlagerung.com/2019/12/06/tactical-theory-set-pieces/
[12] Ojala, M., & Garriga, G. C. (2010). Permutation tests for studying classifier performance. *Journal of Machine Learning Research, 11*, 1833–1863.
[13] U.S. Food and Drug Administration. (2016). *Non-inferiority clinical trials to establish effectiveness: Guidance for industry*.
[14] Chernozhukov, V., Demirer, M., Duflo, E., & Fernández-Val, I. (2023). *Fisher-Schultz lecture: Generic machine learning inference on heterogenous treatment effects in randomized experiments, with an application to immunization in India* (arXiv:1712.04802). arXiv.
[15] Robberechts, P., Van Roy, M., & Davis, J. (2022). *un-xPass: Measuring soccer player's creativity*. StatsBomb Conference.
[16] El-Shesheiny, K. (2025, August 27). Mikel Arteta set-piece rest defence tactics at Arsenal 2024/2025 – How Arsenal stop counterattacks from corners. *Total Football Analysis*. https://totalfootballanalysis.com/set-piece-analysis/arsenal-rest-defence-attacking-corners-set-piece-analysis-tactics
[17] Spearman, W. (2018). *Beyond expected goals*. MIT Sloan Sports Analytics Conference, Boston, MA, United States.
[18] Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine, 15*(2), 155–163. https://doi.org/10.1016/j.jcm.2016.02.012
[19] Mitrotasios, M., Casal, C., Armatas, V., Losada, J., & Maneiro, R. (2021). Analysis of corner kick success in LaLiga Santander 2019/2020. *European Journal of Human Movement, 47*, 8–22. https://doi.org/10.21134/eurjhm.2021.47.2
[20] Litwitz, K., Memmert, D., & Wunderlich, F. (2024). Rebounds in football: A systematic investigation of characteristics of goals scored after rebounded balls in English Premier League seasons 2012/2013 to 2018/2019. *International Journal of Sports Science & Coaching, 19*(6), 2476–2488. https://doi.org/10.1177/17479541241269007
[21] Hudl StatsBomb Data Science Team. (2026). *Set-piece model 2.0* [White paper]. Hudl.

[22] Lamberts, M. (2025, December 31). Introducing expected second ball retention. *Medium*. https://marclamberts.medium.com/introducing-expected-second-ball-retention-adding-value-to-second-balls-1988aa309d9f
[23] Groom, S., Groom, M., Belo, F., Rice, A., Anderson, L., Darvariu, V.-A., & Wang, S. (2026). *Maximising the set-piece return: Optimising football corner tactics with graph reinforcement learning* (arXiv:2606.06353). arXiv.
[24] 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). https://doi.org/10.1145/3292500.3330758
[25] Sportec Solutions & Amazon Web Services. (2022). Bundesliga Match Fact "Set Piece Threat": Evaluating team performance in set pieces on AWS. *AWS Machine Learning Blog*. https://aws.amazon.com/blogs/machine-learning/bundesliga-match-fact-set-piece-threat-evaluating-team-performance-in-set-pieces-on-aws/
[26] Ruiz-Menárguez, G., & Badiella, L. (2026). *The impact of formations on football matches using double machine learning. Is it worth parking the bus?* (arXiv:2602.16830). arXiv.
[27] Cascioli, L., Wang, A., Stradiotti, L., Van Roy, M., Robberechts, P., Wouters, M., Jaspers, A., & Davis, J. (2025). *Quantifying off-ball defensive impact through cover shadows*. Hudl Performance Insights, London, United Kingdom.
[28] Greenhalgh, S., Anghel, D., Kutibashvili, S., Arsénio, H., & Pascual, R. (2025). *Defending the long throw: A dual-model framework for minimising defensive risk*. Hudl Performance Insights, London, United Kingdom.
[29] Bassek, M., Theiner, J., Ewerth, R., Memmert, D., & Raabe, D. (2025). Broadcast analytics – an evaluation of video-based tracking systems with constrained player visibility. *Science and Medicine in Football, 9*(4), 467–477. https://doi.org/10.1080/24733938.2025.2533808
[30] Bates, S., Hastie, T., & Tibshirani, R. (2024). Cross-validation: What does it estimate and how well does it do it? *Journal of the American Statistical Association, 119*(546), 1434–1445. https://doi.org/10.1080/01621459.2023.2197686
[31] Cascioli, L., Goldsmith, M., Stradiotti, L., Van Roy, M., Robberechts, P., Wouters, M., & Davis, J. (2024). *Boot it: A pragmatic alternative to build up play*. StatsBomb Conference.
[32] Guo, C., Pleiss, G., Sun, Y., & Weinberger, K. Q. (2017). On calibration of modern neural networks. In *Proceedings of the 34th International Conference on Machine Learning*.
[33] Błasiok, J., & Nakkiran, P. (2023). *Smooth ECE: Principled reliability diagrams via kernel smoothing* (arXiv:2309.12236). arXiv.
[34] Hopkins, O. (2026, October 24). Why aren't Arsenal scoring any corners? *The Analyst*. https://theanalyst.com/articles/arsenal-corner-struggles-no-goals-premier-league-title

Appendix

A. Model details

Weighting the classes to balance them optimises balanced accuracy at the expense of log loss, costing 0.1594 nats on corners, so it is not used. Calibration has little left to fix: temperature scaling [32], fitted on inner folds only, moves the smoothed calibration error [33] from 0.0093 to 0.0091. Holding out a whole league instead of single matches costs 0.0012 nats, so the model does not lean on the style of a league.

The segmentation is checked against the provider's annotation, and at a fixed expected-goals value a goal is no more likely than another shot to carry a key pass. Over 6,902 shots, 608 of them goals, in 1,692 matches, the difference in that rate is −0.0632 and its interval lies below zero, so the annotation does not favour the shots that scored. The partition rebuilt on the key-pass annotation agrees with the segmentation on every shot after a clearance (Section 2.2 gives the rate over all shots), and the segments separate on the time since the delivery: ranked against the direct shot, time gives an AUC of 0.92 for the flick-on, 0.92 for the prolongation and 0.98 for the shot after a clearance.

Table A1 lists what dropping each block of the design costs, and Section 4.1 discusses the delivery, the only block that excludes zero.

Table A1: Change in log loss when each block of the corner design is dropped (positive means the block helps), with the 95% interval refit inside every match resample, in normal form.

Block	Change in log loss	Refit interval
Delivery	+0.0052	[+0.0023, +0.0082]
Congestion	+0.0001	[−0.0012, +0.0013]
Box depth	+0.0002	[−0.0004, +0.0009]
Counter lane	+0.0001	[−0.0006, +0.0008]
Kick spot	−0.0001	[−0.0008, +0.0005]
Defenders at the edge of the box	0.0000	[−0.0005, +0.0004]
Rebound	−0.0003	[−0.0009, +0.0003]
Marking	+0.0004	[−0.0006, +0.0014]
Defenders per zone	−0.0017	[−0.0035, +0.0001]
Dispersion	−0.0006	[−0.0022, +0.0010]
Attackers per zone	−0.0014	[−0.0035, +0.0007]

The refit intervals are given in normal form because a resample holds about two thirds of the distinct matches, so each draw re-estimates a step on less data, and the cloud of draws, as the

draws themselves show, shifts with the learning curve rather than with the error of the estimate. The normal form keeps only the width the refit adds, and its symmetry was checked: the percentiles re-centred on the point give the normal form's verdict on every contrast that carries a claim.

The subgroup test of Section 4.2 is cross-fitted in each of 20 partitions of the matches. The slope of the per-corner gain on its proxy is 1.12, with p = 0.007, whereas the same test with the proxy's features permuted gives p = 0.97 across rows and 0.543 within matches, so the predictable part of the gain is not an artefact of the cross-fitting.

B. The cells of the box

The eleven cells of Section 2.3 follow the grid that Opta Analyst draws over the box for corners [34]. Inside the box they are its cells: the width of the six-yard box in three equal columns, near, central and far, over two rows, the six-yard box itself and the ground beyond it, and one cell on each side of that width over the depth of the box. Opta Analyst's second row ends at the penalty spot and leaves a central cell from there to the line of the box. Here that cell is split at fifteen yards from goal, its nearer half joining the second row and its further half the edge of the box, which makes the frontal of the three jobs a single cell. Outside the box the cells are the edge, the wide areas on either side and the deep ground further out. The same cells count the attackers and the defenders, so the job of staking the drop, the cell it is read in and the defenders whose effect Section 5.1 sets out to estimate all name one place.

C. Why x ≥ 85

Section 2.1 keeps the free kicks from x ≥ 85, and from that line on a free kick behaves as a corner with offside. The box holds 1.8 attackers at the kick in the band nearest the cutoff and 5.7 in the one nearest the goal line, against 5.7 on a long corner. The median first contact falls inside the box, and the attack wins it on 28.0% and 34.4% of the deliveries in those two bands, against 33.0% on a long corner. Below the cutoff the mechanism changes, not only the distance (Table C1): the first contact lands further from goal and the attack wins it more often, more shots per delivery follow an attacking flick-on or prolongation, and fewer shots per delivery arrive after the first touch. The share after the first touch rises, to 86.4% between x = 60 and 65, only because the direct shot disappears. The cutoff at 85 was chosen after looking at outcomes along the flank, and that selection is declared here.

Table C1: Free kicks taken below x = 85 minus those taken from it, with 95% intervals.

Measure	Difference, pp	95% interval
Share of first contacts won by the attack	+11.8	[+8.8, +14.8]
Shots after a flick-on or a prolongation, per delivery	+1.49	[+0.62, +2.40]
Shots after the first touch, per delivery	−2.90	[−4.66, −1.12]

D. Why eight seconds

The window of Section 2.2 is checked against the possession chain and not borrowed from the literature. The upper panel of Figure 12 reads the shots from the first contact on, second by second, and gives for each second the share of them that still belong to the possession chain of their delivery, those that come from the attack the corner started and not from the next attack of whoever won the ball. The share stays at or above the 95% that a finding needs through the ninth second and falls below it in the tenth, and it falls away after that as the window takes in the other side's attacks. The cutoff sits at eight seconds, a second before the last second that holds, because the second at which a chain breaks is not the same on every play, and a cutoff placed at that last second would count shots that have already left their delivery.

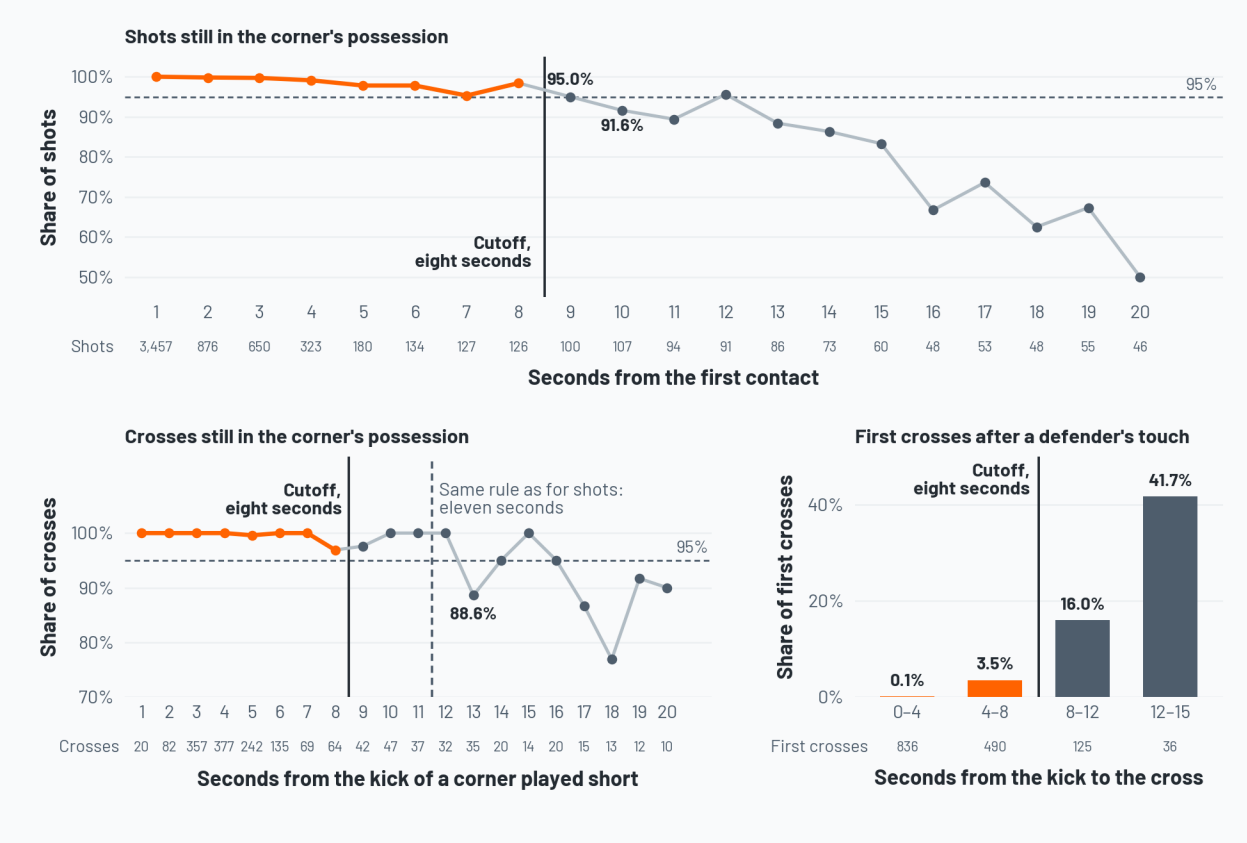


Figure 12: Shots (above) and short-corner crosses (below left) still in the corner's possession, and first crosses after a defender's touch (below right).

For a corner played short, counted from the corner kick, the share of the side's crosses that still belong to the possession chain of the corner stays at or above 95% through the twelfth second and falls below it in the thirteenth (lower left panel of Figure 12). The rule that fixed the eight seconds of the episode, a second before the last second that holds, would therefore put the wait for the cross at eleven. The chain alone does not tell whether the routine was interrupted on the way, and the longer a cross takes, the more time a defender has to touch the ball first. Within eight seconds of the corner kick fewer than one in twenty of the first crosses follow a touch by a defender, and

between the eighth and the twelfth second more than one in ten do (lower right panel). Hence, the wait is the same eight seconds as the episode, and not eleven.