Beyond the First Defender in Football Pressing

A causal framework for evaluating pressing efficiency using broadcast tracking data.

Jainil Kakka · Brown University · Hudl Performance Insights 2026

Abstract

Pressing is how a defending team tries to win the ball back, or to stop the team that has it from progressing. It has been coached in vastly different ways over 50 years, but almost every way of measuring it counts what the defender nearest to the ball does. This paper aims to ascertain whether those actions change what happens next. Across 30 Serie A matches from 2025/26, with event data and tracking at 25 frames per second, we detect 15,203 presses from the defending player actions towards the ball. For each pressing event we measure what the nearest defender did, what their teammates did, and what followed over the next 5 seconds as the outcome. The difficulty is that defenders do not choose at random. They close fast on a slow, isolated carrier and hold off a set one, so comparing fast closes with slow ones compares different situations. We use double machine learning to hold the situation fixed, such as where every player is and how every player is moving. When situation is held fixed, nothing the nearest defender does changes the outcome, not their closing speed, their approach angle, whether they block the forward pass, whether they arrive front-on, or how many go with them. What everyone else does changes the outcome, and the actions that halt an attack are not the actions that win the ball. An average press allows 5.7 m progression. Two actions cut that number significantly. Stepping the defensive line up as the press goes in reduces progression by 5.5 m, and more defenders converging towards the ball a further 2.6 m. Together about 8% of the pitch, denied in 5 seconds, but neither wins possession back. The second defender closing hard (+1.3 turnovers per 100 presses) and markers staying tight on the nearest passing option (+1.2) are the two actions that turn presses into turnovers. Tight marking the nearest passing option is the only action that suppresses progression and wins possession at the same time. This is close to what the modern game already assumes. Pressing has evolved from a team chase into a structured shape that triggers on a cue, and the numbers follow the shape rather than the chase.

1 Introduction

1.1 Evolution of pressing

Pressing, as we know it, began in 1934 with hockey coach Thomas Patrick Gorman. He told his players to challenge opponents near their own goal and block passing lanes, which led to initial losses but eventually helped them dominate the National Hockey League. In the 1970s, Dutch football teams like Ajax adopted a similar approach, which we today known as "Total Football". The Dutch Manager, Michels said "*For me, it would be better to call my game 'pressing football'. This is what we wanted to create with our Ajax side and the Dutch national team in 1974: create a basic*

game where all 10 outfield players push forward even when we don't have the ball. We're always pressing forward". This strategy involves all players pressing forward even without the ball. Today, terms like high block, low block, and Gegenpressing describe different ways teams use pressing to quickly regain possession of the ball.

What is most common now is different again. Many sides hold a compact mid or low block, concede the ball in areas that do not threaten them, and press only when the ball enters a zone chosen in advance. Simeone's Atlético Madrid is the clearest case: two narrow banks of four, the ball pushed wide, numerical superiority created in the ball's zone, and the opponent forced backwards rather than dispossessed. In 2015/16 they conceded 18 goals in 38 league games. The aim of that press is often not to win the ball at all. It is to stop the opponent advancing into the part of the pitch where advancing hurts. Defensive styles differ systematically between teams and leagues along exactly these lines (Fernandez-Navarro et al., 2016).

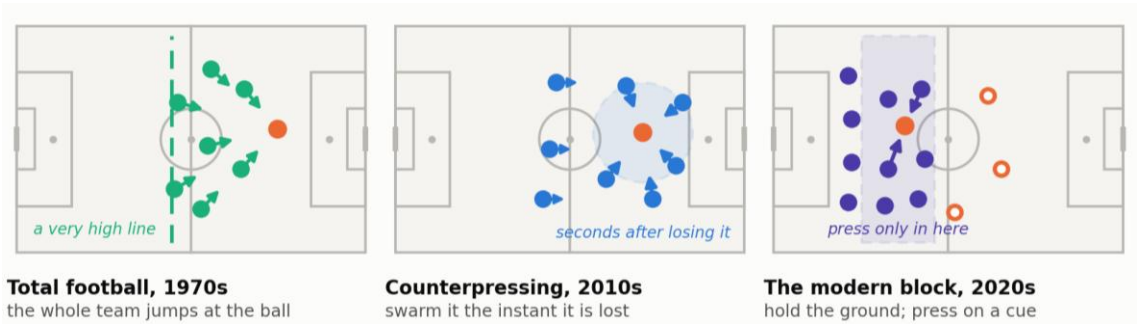


Figure 1. Evolution of Pressing.

Three things are worth drawing out before any data is introduced. First, in none of these eras is pressing one player's job. Second, what counts as a successful press is not the same in all three: the 1970s press and the counter press are trying to win the ball, while the modern block press is often trying to deny ground. Third, the defending team's shape and movement are the instrument in every case, and the player nearest the ball is the visible part of something larger. That third point is the motivation for this paper.

1.2 How pressing is defined

Academic definitions have tried to capture this. Pressing is described as defensive pressure applied to the ball carrier and nearby passing options with the aim of forcing an error or slowing play (Fernandez-Navarro et al., 2016). Others define it through its spatial effect such as a reduction in the time and space available to the player on the ball (Link et al., 2016). Bauer and Anzer (2021) define a press as a coordinated movement of defenders toward the ball carrier detectable from tracking data. Each definition captures something real, but none captures everything. A press can be one defender sprinting alone or eight moving in unison. It can last half a second or five.

1.3 Outcome of a successful press

Because the eras disagree about what a press is for, a study of pressing has to say what it will count as success and say it before looking at any result. We use two measures, and we use both from the start.

- **Ground denied.** How far does the ball travel upfield in the five seconds after the press begins? This is the currency of the modern block press — territory denied, regardless of whether the ball changes hands.
- **Possession won.** Does the ball change hands within that same window? This is the currency of the 1970s press and the counter press.

Keeping them apart is the single most consequential decision in this paper. A press that concedes nothing and never wins the ball has done one job well. A press that wins the ball back but is expensive when it fails has done the other. Collapsing both into one number hides which job a team is doing.

2 Data

2.1 Source

The data is proprietary, supplied by Hudl for the Performance Insights Research Competition. It covers 30 Serie A matches from the 2025/26 season involving six clubs: AC Milan, AS Roma, Como, Inter Milan, Juventus and Napoli. Every match carries two synchronized feeds.

- **StatsBomb 360 event data.** Every on-ball action with its type, outcome, possession identifier and a freeze frame of player positions. The schema follows the conventions now standard in public football datasets.
- **Positional tracking.** Derived from broadcast footage at 25 frames per second, carrying position and velocity for all 22 players and the ball. Coordinates are in metres on a 105 by 68 m pitch with the centre spot at the origin, so x runs from –52.5 to +52.5 and y from –34 to +34. Appendix C lists the full set of conventions.

The two feeds share a clock, so every event can be placed in a tracking frame and every frame attributed to a possession.

2.2 Press detection

Instead of using the event feed's pressure annotations to define a press. We detect presses from the movement itself, using three conditions that must hold together.

- The defender is within 15 m of the ball carrier.
- The defender is closing on the carrier at 2.0 m/s or more.
- Both of those holds for at least four tenths of a second.

A 15 m zone-dependency and a speed check keep one sustained approach from being counted several times.

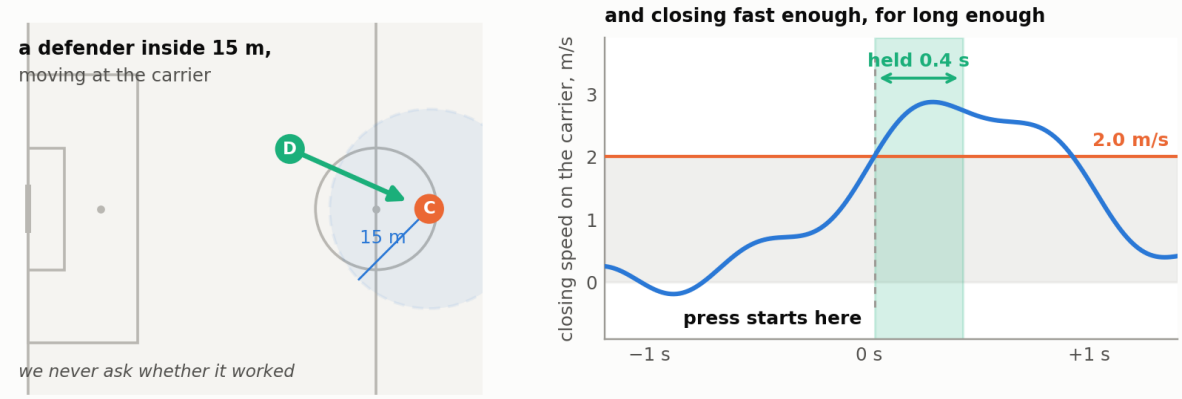


Figure 2. Detecting a press.

Our method flags 25,830 presses across the 30 matches, about 861 a match. In comparison the StatsBomb hand-coded events carry 10,451 such cases in the event feed, about 348 a match. A hand-coded Pressure event is logged by a human when a defender's closeness visibly affects what the player on the ball does. A movement-based rule records every committed approach, whether the attacker reacted or not. For a causal study the second is what is needed, because selecting presses on whether they visibly worked means selecting on a consequence of the very thing being measured causing selection bias.

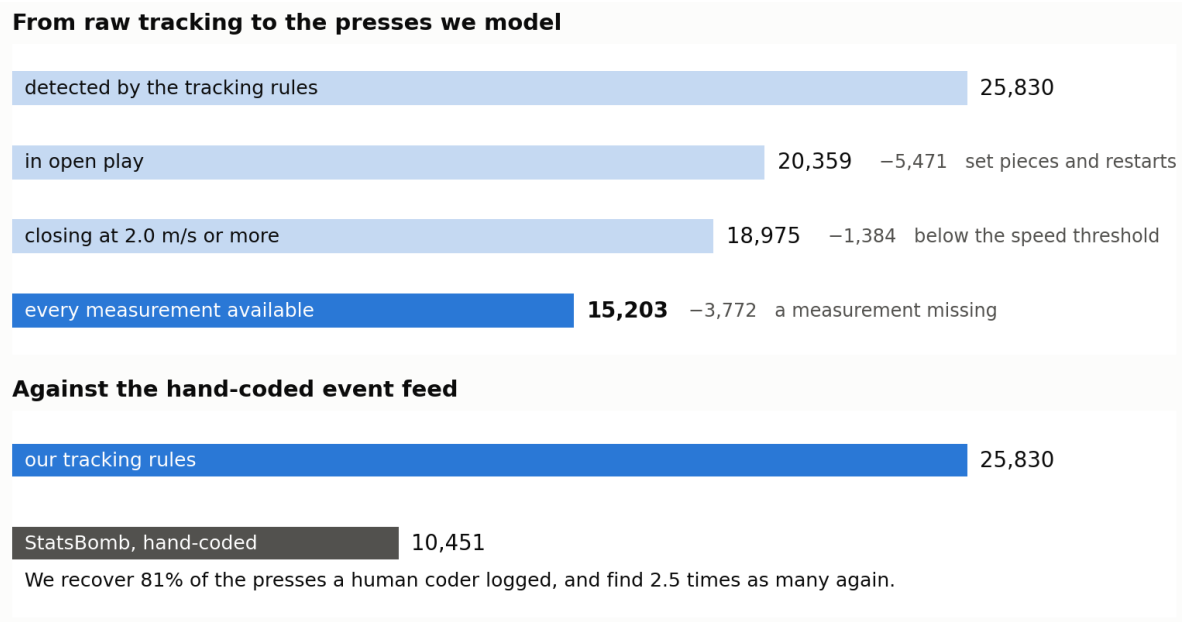


Figure 3. Pressing event sources.

Four in five of the hand-coded Pressure events have one of our detections within two seconds, which is a reasonable check that the rules are finding real pressing rather than noise. Of the 25,830, we keep open play only, dropping 5,471 presses that follow a set piece or dead-ball restart, where the arrangement of players is not chosen in the way the rest of the study assumes. A further 1,384 records fall below the speed threshold once measured precisely. That leaves 18,975. Of those, 15,203 have every measurement the first model needs.

2.3 Feature engineering

We calculate three kinds of things. What the player nearest the ball does, labelled A. What their teammates do, labelled B. And what follows, labelled Y. Everything else about the moment is context which we try and control for, and section 3 explains what we do with it.

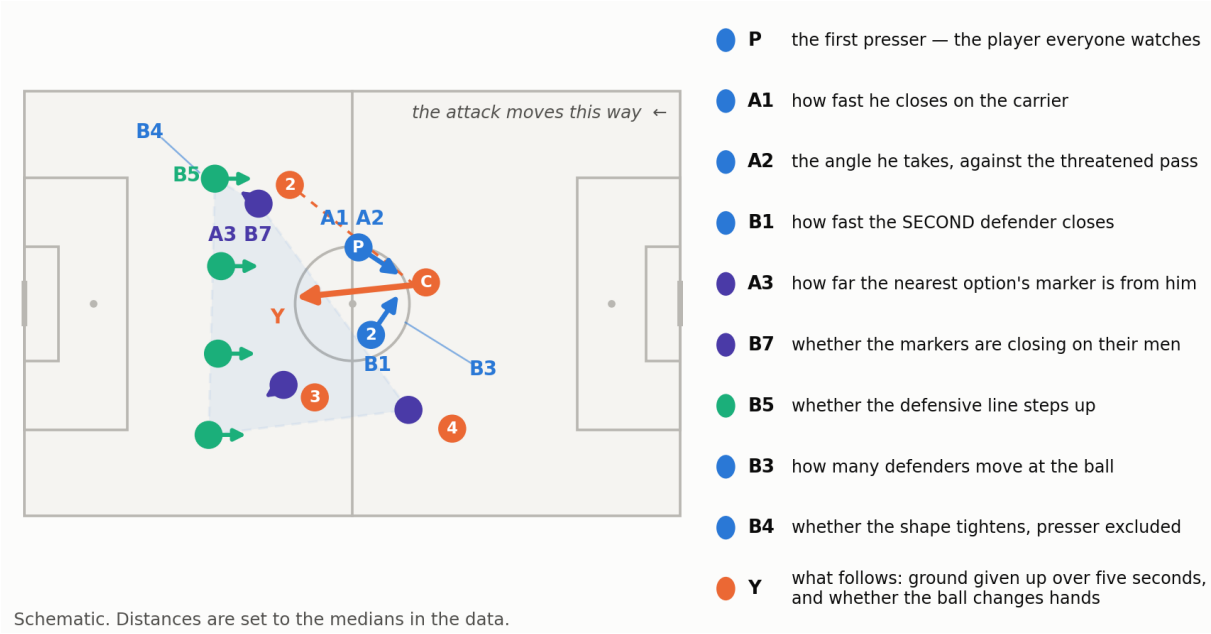


Figure 4. Features present in the study.

2.4 The outcome variables

Section 1.3 sets out the two kinds of success a press can have. Each is measured over the same five-second window that begins when the press is detected.

- **Ground conceded** is the distance the ball travels upfield between the start of the press and five seconds later, measured along the attacking direction. The clock stops early if possession changes hands, and the ball's position is frozen at that frame.
- **Ball won** is simply whether possession changed hands in the same window. It happens on 18.2% of presses (roughly one in five).
- **Three-level scale.** One variable that captures both kinds of success together. 3,447 presses end with the attack losing the ball, 9,843 presses end with neither side gaining a clear advantage and 5,685 presses end with the attack making 10 m or more. This is the variable that lets us ask whether the press did one job, both or neither.
- **On-ball value.** Built from the family of frameworks that score each action by the change in scoring chance it produces. The event feed carries a value for every action we sum it across the attacking team's actions inside the same window, so it measures the scoring threat the attack generated while under pressure. It weighs a 10 m gain in a dangerous area above 10 m gain in a harmless one. (We do not report results for this measure. At the level of a single five-second window almost every press produces no on-ball value at all. The

distribution is a spike at zero with thin tails, and our context model predicts worse than the outcome's own average at every specification we tried.)

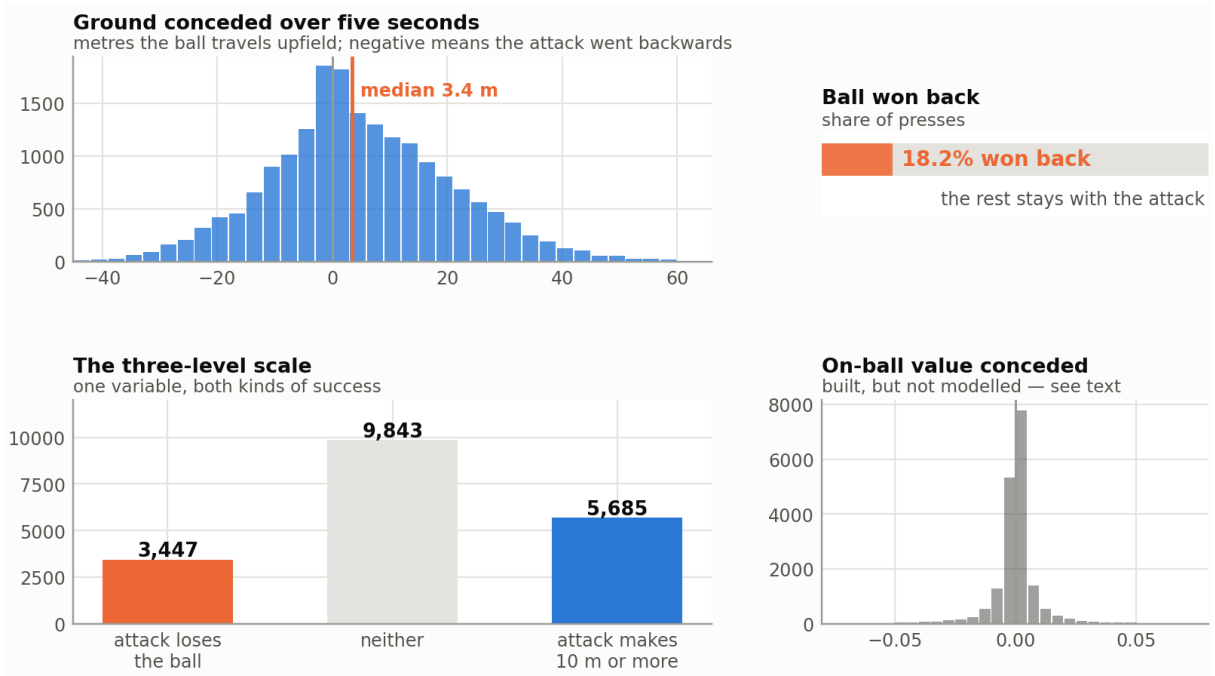


Figure 5. The outcome variables descriptive statistics.

Keeping ground and possession apart matters more than it sounds. Presses that win the ball back and presses that do not concede almost the same ground on average, 4.95 m against 4.87 m, because the ball travels faster in seconds before a turnover and the window is shorter. A single measure of ground conceded therefore cannot tell the two apart, which is why both outcomes are carried through every result in the paper.

3 Method

The contribution is in three parts. A press detector (Section 2.2) that works from movement rather than from annotation, so that presses are not selected on whether they visibly worked. An estimation design that holds the situation fixed using both the positions of all 22 players and their movement, with a reported diagnostic for every action saying whether the question could be answered at all. And a result, reported against both measures of success, that locates pressing effectiveness behind the ball rather than at it.

3.1 Estimating, not predicting

Most papers in this space build models that predict the future or the likelihood of a future event based on the current circumstances. That model would maybe predict the future better than anything in this paper, its accuracy comes from the whole situation at once and it has no way to separate the actions the defender opted for from the actual situation they were handed.

The core problem is that the defender's actions are not independent of the situation, situations dictate how a player reacts and performs. Since the actions and situations are not independent, using them both to predict the outcome leads to inferences that are riddled with confounder bias. Think of it like this, imagine we want to know whether closing hard actually causes turnovers, or whether turnovers just tend to happen in situations where defenders close hard anyway. The only way to find out is to strip the situation out of both sides of the questions. To get around this we use double machine learning.

This is why we set up the smaller questions first, which is also the older discipline. It costs accuracy and buys interpretation, it produces a number (effect) per action (intervention) in meters or in turnovers, with an interval around it.

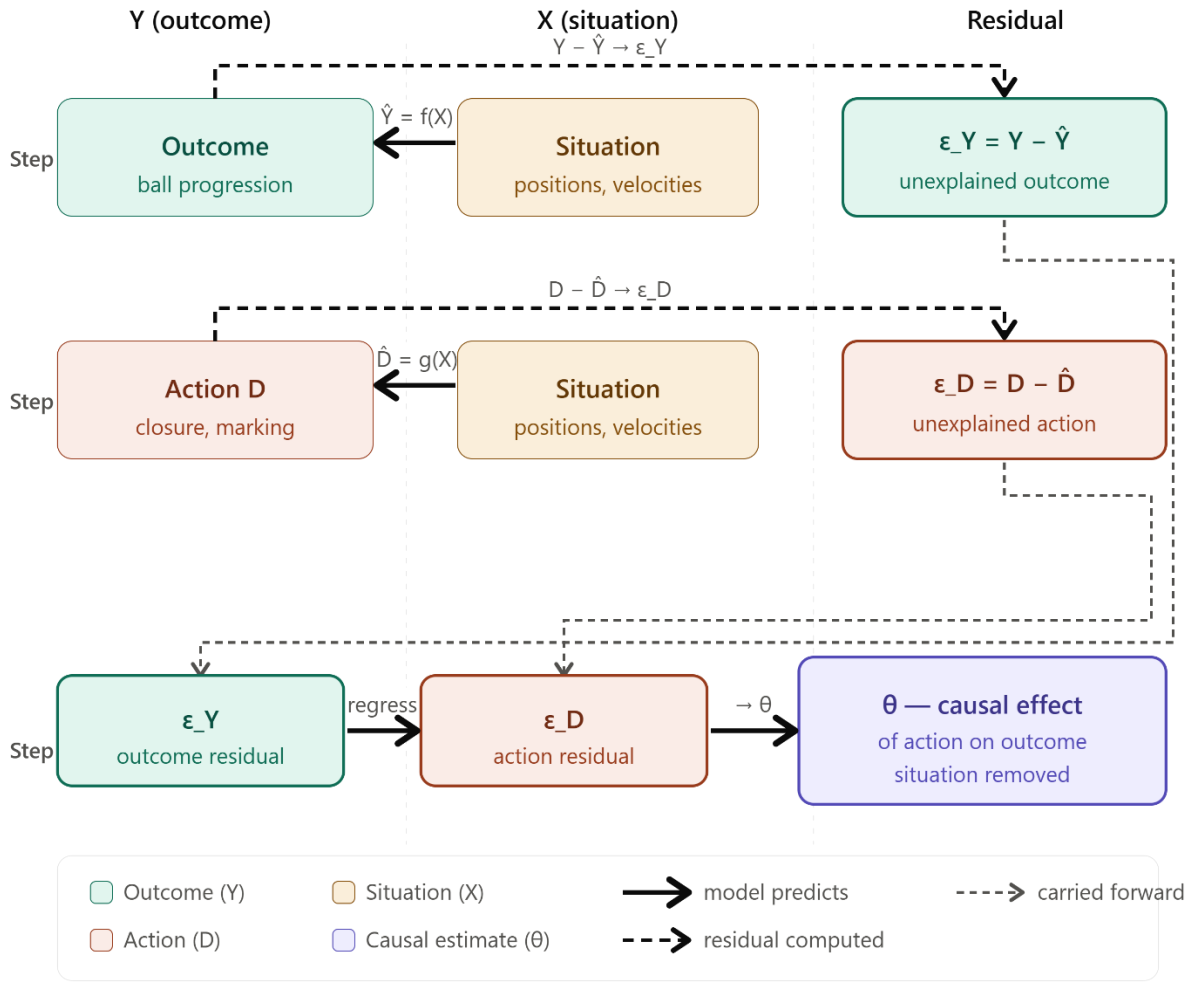


Figure 6. Double machine learning.

First, we take the outcome such as how far the ball progressed (in 5s) and ask how much of that a machine learning model can predict from the situation alone. Whatever the model cannot predict, the part left over, is the outcome that the situation does not explain.

Second, we do the same thing to the action. We ask how hard a defender would be expected to close, given that same situation. Whatever is left over is the part of the action that the situation did not force, the part that reflects a genuine choice.

Third, we relate one leftover to the other. Does a defender who closed harder than the situation demanded produce a better or worse outcome than expected? That relationship between the leftover action and the leftover outcome is the causal estimate. Because both sides have had the situation removed, what remains is the effect of the choice itself rather than the circumstances it was made in.

3.2 Modelling safeguards

If we train the model on all 15,203 presses and then ask it to predict those same presses, it has already learned the quirks of each individual press during training. Its prediction quietly absorbs some of the defender's actual choice rather than just the situation. We subtract part of the very thing we are trying to measure. The residuals are no longer truly unexplained. They carry traces of the action inside them and the estimate shrinks toward zero not because the action has no effect but because the method ate its own answer.

The fix is to never let a model predict a press it was trained on. We divide the 30 matches into five groups and predict the presses in each group using a model trained only on the other four. Every press is predicted by a model that has not seen it. This is also called cross-fitting.

3.3 Standard errors

The DML method in section 3.1 gives us an estimate of the causal effect. But an estimate alone is not enough, we also need to know how much to trust it. Presses within a match are not independent of each other. They share a game state, a referee, a pitch and two tactical plans, so treating them as 15,203 independent observations would overstate how much evidence there is. They are similar than two presses from different matches. Due to which confidence intervals end up too narrow meaning we would call things significant that are not.

Standard errors are therefore clustered by match, which treats each match rather than each press as the independent unit. With 30 matches, that is a small number of clusters, and the usual cluster-robust formula is known to be too narrow in that regime. We correct for this using a wild cluster bootstrap, which simulates the uncertainty 999 times under the assumption of no effect and builds the confidence intervals from those simulations. The result is a deliberately conservative test, where ordinary standard errors would be far narrower here and would make several of our nulls look like findings. When a test clears the bar in this paper it clears a high one.

3.4 Cases of null residuals

The DML method works by removing the situation from the action and estimating the effect of what is left, the part that reflects a genuine choice. But some actions are almost entirely dictated by the situation. If defenders always respond the same way in the same circumstances, there is no variation left after the situation is removed and no choice to measure. The method will still

produce a number, but that number is built on noise. A null result in that case is not a finding, it just means there was nothing to work with.

To guard against this we check how much of each action the situation failed to explain and report that leftover variation for every action in the paper. We set a floor at 10%, below it an action is treated as unanswerable rather than as null because those are two different claims and only one of them is honest. On constructed data where the true answer was known the estimator missed it entirely with 7% variation left and recovered it correctly with 34% left. This matters because most of our results are nulls and a null is only worth reporting when the method has a genuine chance of finding something.

4 Results

4.1 The trigger

We began with the five things that player controls. We name these as A prefix and have 5 of them given in the table below.

	What it measures
A1	How fast they close on the carrier
A2	The angle they take, relative to the pass the carrier most threatens
A3	How far the nearest passing option is from the defender marking him
A4	How compact the pressing team's shape is
A5	How many defenders are pressing at once

These build on a long line of spatial descriptions of football, from dominant-region tessellations (Taki and Hasegawa, 2000) through danger models built on tracking (Link et al., 2016) to contextual descriptions of team shape (Bauer et al., 2023). All five enter together, so each number is the effect of that action with the other four held fixed.

That picture of the moment (or the situation as we call it) is essentially what a freeze frame gives us, and it is the level of detail most published pressing analysis works on. A freeze frame cannot say how fast anyone is moving or in which direction, and it cannot say whether a shape is tightening or spreading, because those need two frames and preferably fifty. So, we ran the same analysis three times, each time with more of what the tracking data offers, and it explained below.

Picture of the moment	What it adds	Needs
A snapshot	Where everyone is: ball, distances, space, defender counts, score, minute, workload	a freeze frame
+ movement	How fast the ball, the carrier and each team are moving, and whether the defending shape is tightening	continuous tracking
+ pitch context	Depth of the defensive line, distance from ball to that line, how spread it is, distance to goal, who is goal-side	continuous tracking

If we only consider the snapshot data or freeze frame data, the angle the presser takes is worth 5.25 m of prevented progression for a shift from a typical shallow angle to a typical head-on one ($p = 0.001$). Closing speed does nothing measurable ($p = 0.121$), nor does shape or the number of pressers. But tracking data opens a number of possibilities such as velocity and pitch context in terms of the position to other players and their directions towards each other.

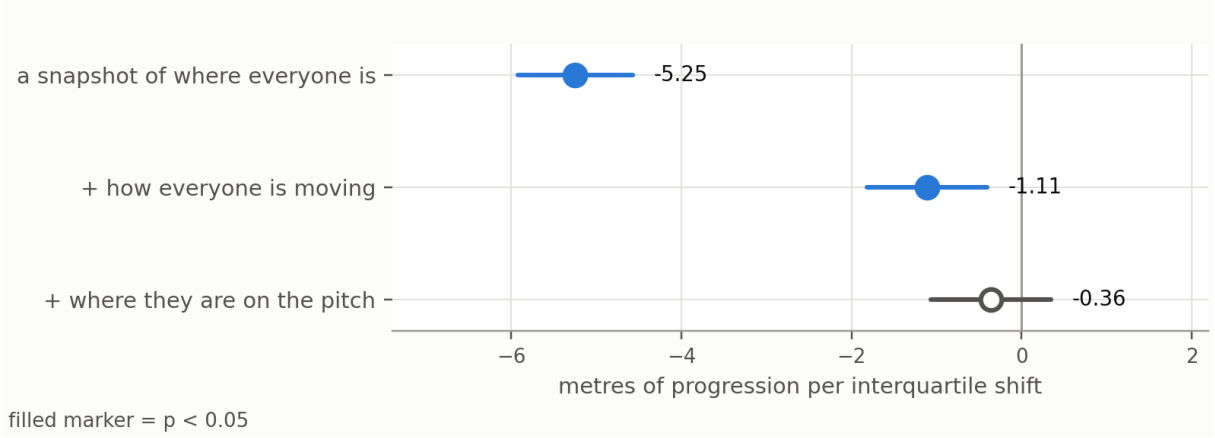


Figure 7. How data quality affects results.

The presser's run direction goes from the largest effect on the pitch to nothing at all. Worth 5.25 m against a snapshot, 1.11 m once movement is included, and 0.36 m with the whole picture, at which point it is indistinguishable from zero ($p = 0.311$).

The change in result is not because there is no variation left for the model to predict from, in fact the variation only changes from 0.58 to 0.38. This means when taking the broader situation in view the action does not affect the result and the model's predictive accuracy improves steadily from 18% to 26% to 29%, so the extra information is earning its place rather than soaking up noise.

The reason is easy to understand; a defender's run direction correlates at –0.68 with their own team's collective movement and –0.61 with the carrier's. It is largely just the direction the whole passage of play is flowing. Once that flow is measured the individual adds almost nothing on top.

Two readings a tactical audience would be interested in have the same result. Blocking the passing lane is worth 0.09 m ($p = 0.670$) and arriving front-on rather than side-on is worth 0.21 m ($p = 0.302$). Against the second outcome (whether the ball changes hands), the closing speed returns $p = 0.540$ and run direction $p = 0.084$, changing sign between windows and carrying the least variation left of anything we measure. So, we do not consider the effect.

4.2 Team response

What we understand from the results from the A (presser movement) set of variables is that, if the first defender's apparent effect belongs to the team's collective movement, then that movement is where the effect lives. A block's forward velocity, how it compacts, where its line sits: these are coached, drilled and chosen. So, we measured them directly.

This introduces a problem that was not present in the A (presser movement) analysis. A compact block is compact for reasons that are subject to everything else about the situation or moment, so compactness as a static state cannot be treated as an action. What can be treated as an action is the change: whether the block compacted in response to this press.

We therefore split the moment in time. The defending team's configuration one second before the trigger goes into the context as a static fact about the moment. The change from there over the next second becomes the action. We change the question from whether the team was compact, to whether they compacted in response, which is also the question a coach can act on, since the static arrangement is the plan and the change is the execution.

Seven actions enter together, with the first presser's own actions held as controls, such that nothing from the B (team movement) rides on the effects of the A (presser movement) set of variables.

	What it measures
B1	How fast the second defender closes on the ball
B2	Whether a second defender properly engages at all, and how many do
B3	How many defenders are moving towards the ball
B4	Whether the defending shape tightens, with the presser left out of it
B5	Whether the defensive line pushes up
B6	How many of the carrier's options are covered by someone other than the presser
B7	Whether the markers are closing on the men they are covering

B4 and B6 were written down as the two we expected to matter before any of these were estimated, on the reasoning that squeezing the shape and covering the options are the two mechanisms a coach would ideally name. We report below what became of that expectation, because a reader is entitled to know which results were predicted and which were found.

4.3 Keeping the attack out

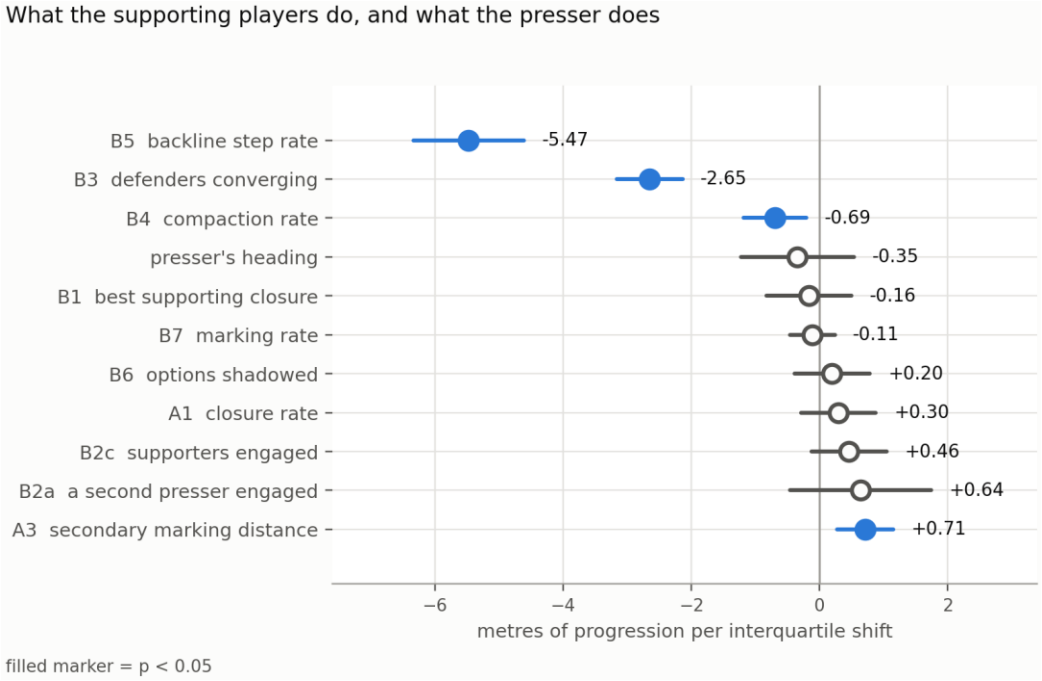


Figure 8. A, B Progression allowed.

Two things stop an attack making ground, and both describe the shape behind the ball and not the duel at it. The defensive line pushing up as the press goes in is the largest effect in the study. Moving from a line that is dropping to one that is advancing, a swing of about 2.2 m/s, which is the difference between a typical retreat and a typical step, is worth 5.47 m of prevented progression (95% interval 4.61 to 6.34 m), about a third of a standard deviation. Defenders converging on the ball are worth a further 2.65 m (2.14 to 3.17). Both are around twice as large in the second after the trigger as in the second before it. That pattern is what the two windows were built to detect: a block that happened to be moving already would show the same effect in both, while a block reacting to the press shows more after it.

4.4 Winning the ball back

Run the same set of actions (A, B) against whether possession changed hands and a different set of names comes up.

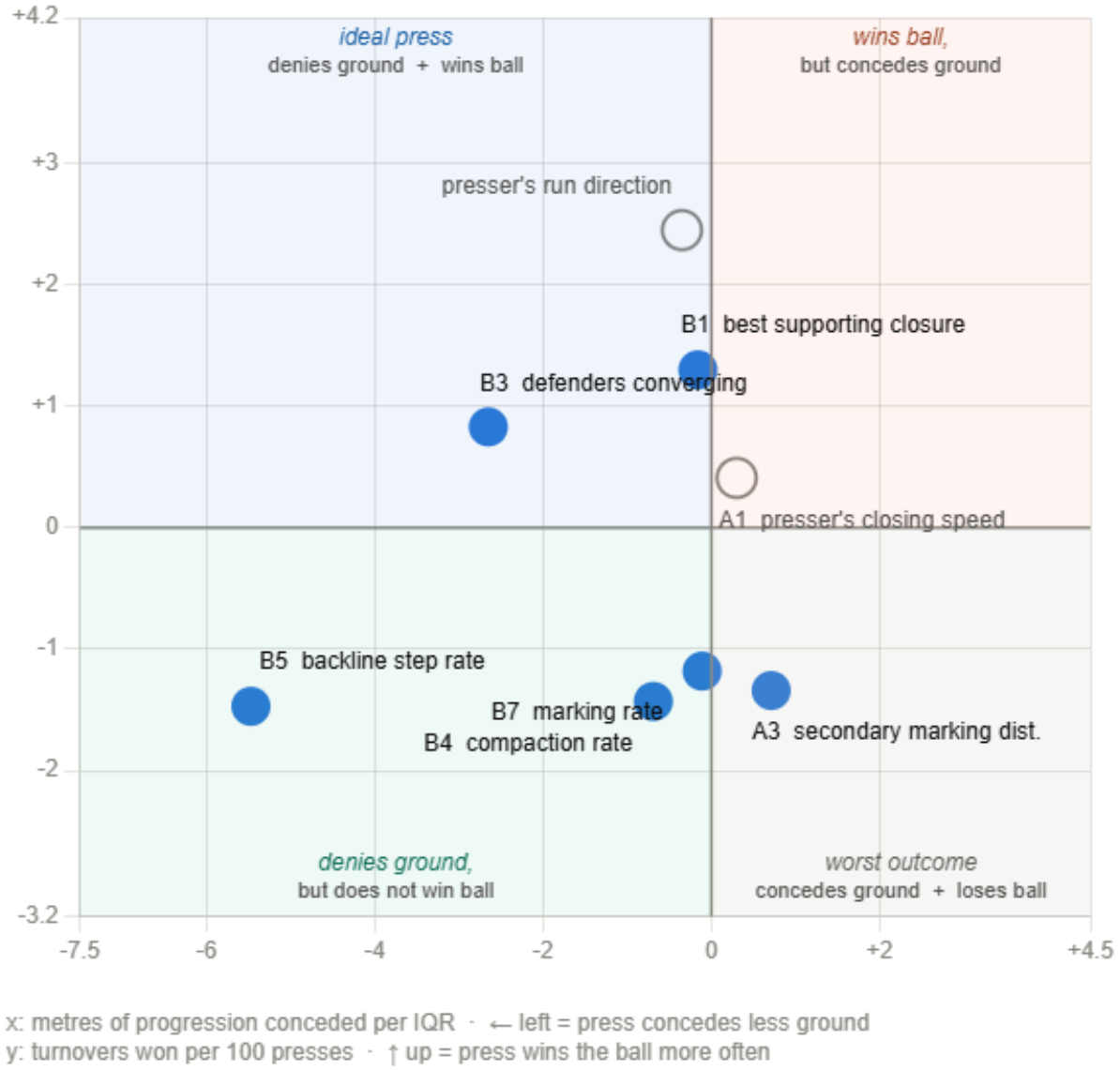


Figure 9. Action scatter plot over effects on the outcome variables.

In this chart the filled or blue dots show variables that have a significant p-value, provided in detail in table below. The scatter plot separates the actions cleanly into two groups. The actions that stop the attack making ground such as the line stepping up and defenders converging but they do not win the ball back more often than average. The actions that win the ball back are the second defender closing hard and markers tightening but they do nothing measurable to ground conceded. The two jobs are done by different players.

Action	Ground conceded (m)	p	Ball won (pp)	p
B5 line pushes up	**–5.47**	0.001	–1.47	0.080
B3 defenders converge	**–2.65**	0.001	+0.83	0.276
B1 second defender closes	–0.16	0.630	**+1.30**	**0.025**
B7 markers tighten	–0.11	0.526	**–1.18**	**0.011**
A3 marking distance	**+0.71**	**0.003**	**–1.34**	**0.043**
B4 shape tightens	–0.69	0.011	–1.43	0.013
presser's run direction	–0.35	0.461	+2.45	0.084
A1 presser's closing speed	+0.30	0.324	+0.41	0.540

Tighter marking on the nearest passing option reduces both the ground conceded and increases the chance of winning the ball back. Two independently built measurements such as the distance the marker ends at and the rate at which they close it which points the same way.

The second defender's effect on turnovers is the most robust finding in the paper. It holds across both time windows at 1.44 percentage points before the press trigger and 1.30 after.

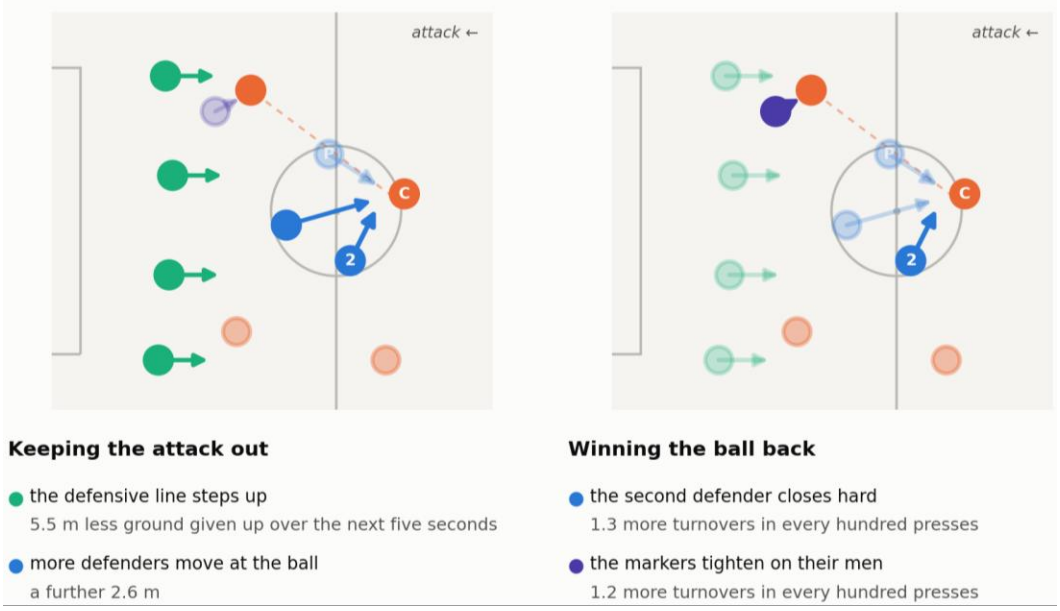


Figure 10. Result comparison between outcome variables.

4.5 Actions with no measurable effect

Four actions we measured return values which indicate close to no effect, and in a study of this kind that is information rather than an absence of it. Each of the four had enough variation left after the situation was accounted for that an effect could have shown up.

Covering the carrier's passing options, measured as the share of the options shadowed by someone other than the presser, is worth 0.20 m ($p = 0.484$). This was one of the two we named in advance, so this is a prediction that did not survive contact with the data. Models of pass completion treat covered options as a first-order determinant of whether a pass is available at all (Anzer and Bauer, 2022), and our reading is that this is real but already priced in: the options are covered or not covered a second before the press begins, which puts it in the static arrangement rather than in the response.

Whether a second defender engages at all is also null ($p = 0.380$). There is almost no timing left to recover here: once a defender qualifies as a second presser, they arrive at a median 0.08 s after the first, because the detector's own conditions mean they qualify almost at the same instant. The variable is close to a constant, and a near-constant cannot explain variation in anything.

How many support pressers arrive is null too ($p = 0.607$), which sits oddly beside the finding that more defenders converging on the ball prevents 2.65 m. Converging counts every defender moving towards the ball, including those who never arrive and never contest it; engaging counts only those who meet the full detector conditions. The reading we take from the pair is that the value is in the collective movement towards the ball rather than in the number of bodies that end up in the duel, which is the same conclusion as section 4.1, measured from the other end.

Squeezing the shape, the other action we named in advance, does reach the usual threshold at 0.69 m ($p = 0.011$), but it does not survive the check in section 5.1. We report it as suggestive rather than as a result.

5 Robustness

5.1 Is the line pushing up only when the attack is already retreating?

This is the obvious objection to the largest result, and it needs to be tested to make this claim. A defensive line advances when the attack turns back, and an attack turning back was not going to make ground anyway. The data shows that the line's movement correlates about −0.57 with how far the ball travelled in the seconds beforehand.

So, we control it in two ways, we first add the ball's movement over the second strictly before the action (That is part of the situation and without a doubt right to account for), and the second adds the ball's movement over the same second as the action. If pushing the defensive line up is itself what sends the ball backwards, then controlling for it removes part of the effect being measured. The results are shown below.

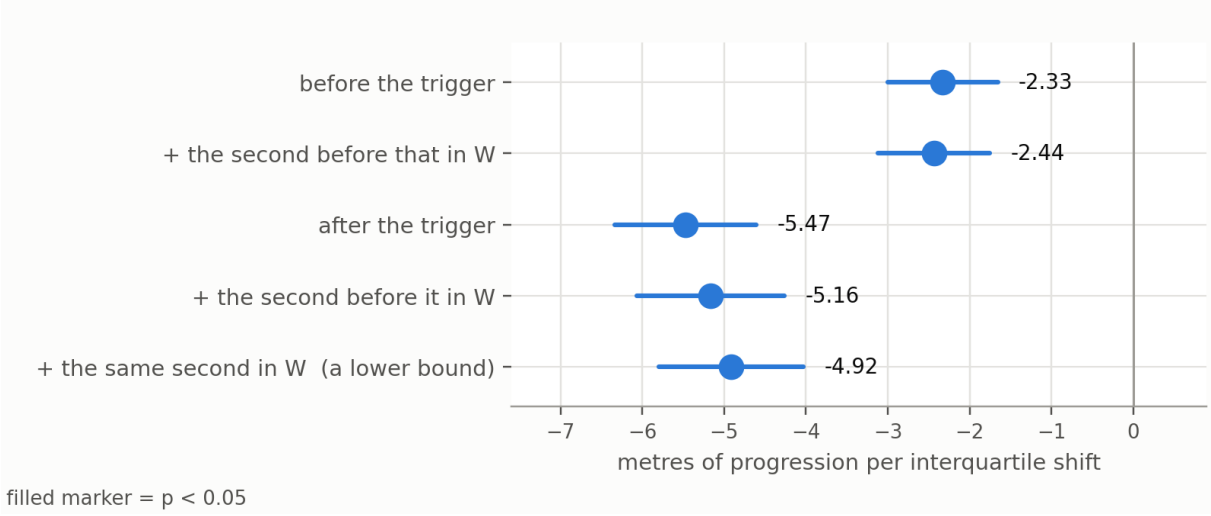


Figure 11. The line pushing up, under every control for the attack's own retreat. The effect survives all five.

The effect loses 5% under fair control and 10% under the harsh one and stays clearly distinguishable from zero in all five versions whereas the pre-trigger estimate grows slightly under its own control. Defenders converging behave the same way, going from 2.65 m to 2.46 and 2.49.

5.2 Could the method be producing effects that are not there?

A method that manufactures findings would manufacture them from noise as readily as from signal, so we gave it noise and checked. We took the real data and jumbled the outcomes within each match, so that each press was paired with the outcome of a different press from the same game. This destroys any genuine relationship between an action and what followed it, while leaving everything else such as the actions, the situations, the sample sizes, the clustering the same. Any effect the method reports on this data is an effect it invented.

It reported essentially nothing, which is the right answer. Comparing the real estimates against the spread of what the shuffled data produced, the line pushing up sits more than four times outside that spread and every action we report as null sits inside it.

5.3 Does the answer depend on how success is defined?

Ground and possession are reported separately because they are different things. A reader may reasonably prefer them combined, and we should not be able to produce our conclusions only by choosing to split them. So, we also scored each press on the single three-level scale from section 2.4. The attack losing the ball at one end, the attack making ground at the other, and everything else in between. The threshold for "making ground" is an arbitrary choice, so rather than fixing it we ran the whole analysis at 5 m, 10 m and 15 m.

The conclusions hold at all three, the line pushing up, defenders converging and marking distance stay clearly distinguishable from zero. The presser's run direction stays indistinguishable from it. We treat this as a check rather than a headline, because putting both outcomes on one scale assumes that the step from losing the ball to containing the attack is the same size as the step from containment to progression, and that is an assumption rather than something we measured.

We also varied the detection threshold itself, since 2.0 m/s is a choice. The direction of every reported effect is stable across the sweep.

5.4 Multiple comparisons

Eleven actions per window at the usual threshold would produce about half a false positive by chance, and we run more than one window. Protecting against reading noise as a result is not a correction factor but the pattern of where an effect shows up.

We named two actions in advance and report plainly that one did not hold and the other was null. The line pushing up and defenders converging are reported as results because they hold at p = 0.001 across five specifications, two outcome windows and three different pictures of the situation, not because they cleared a threshold once. The possession results, as section 4.4 says, are weaker on this test and are labelled exploratory for that reason.

6 Discussion

6.1 Why the first man's null is what the modern game predicts

The result most likely to be quoted from this paper is the one in section 4.1, nothing the player nearest the ball does changes what happens next. It deserves more than a line, because read carelessly it says something false, and read carefully it says something the game already believes.

What it does not say is that the press does not work, or that the first defender's job does not matter. **What the result says is that among presses that have already happened, the variation in how the nearest defender executed their part does not explain the variation in what followed**.

The modern block press works because ten players hold a shape and the cue to step is collective. In every case the first defender is the visible edge of a coordinated movement, and in every case the coordination is the mechanism.

A coach reviews a conceded goal, freezes the frame the moment the press begins, and observes that the first defender approaches from the wrong side. The observation is usually correct, the inference that a better approach would have changed the outcome is the part our data does not support, because in the current frame their approach was already being determined by where their teammates were and which way they were going.

6.2 Two jobs, and what they say about the low block

Section 4 separates two jobs that the pressing literature usually treats as one. Keeping the attack out is done by the line stepping up and by bodies converging on the ball. Winning the ball back is done by the second defender closing hard and by the markers tightening.

This basically shows us that a low or mid-block that presses on a trigger is optimizing the first job of denying ground, and the two actions that deny ground are exactly how the low block is setup as where they step up as a unit and bodies move towards the ball together rather than individually. Our largest effect is a measurement of what a block is for. The line stepping up is worth 5.47 m of prevented progression, which is roughly a third of a standard deviation and larger than anything

else we measured. A team that holds its shape and steps together is buying territory, and the amount it buys is now a number rather than a conviction.

The practical reading for a side that plays this way is that the two jobs need separate instructions and separate review. If the aim on a given night is territory, the things to coach and to grade are the line's step and the collective movement towards the ball, and a low tackle count is not evidence of failure. If the aim is to win the ball back in a chosen zone, the instructions are different ones to different players: the second defender has to close hard, and the markers must tighten. Asking for both at once is possible, but with this evidence only one action serves both aims.

6.3 Breaking a block

If the results describe what makes a block work, they also say something about attacking one.

If we invert the defensive results into attacking one, we get two ways or methods. The first is that the attacking actions worth most are the ones that stop the line stepping as a unit, because the step is where the territory is bought. Anything that makes the line hesitate like a runner in behind, a player occupying the space the line wants to advance into, helps in stopping the line advancing.

The second concerns the one action that helps the defending side on both counts. Marking the nearest passing option tightly reduces ground conceded and increases turnover. Its mirror image is that the single most valuable attacking movement near a press may be the one that frees the nearest option rather than the one that helps the carrier. That is a hypothesis our design can state but not test, and section 7 lists it as the first thing we would want to measure next.

6.4 On the training ground

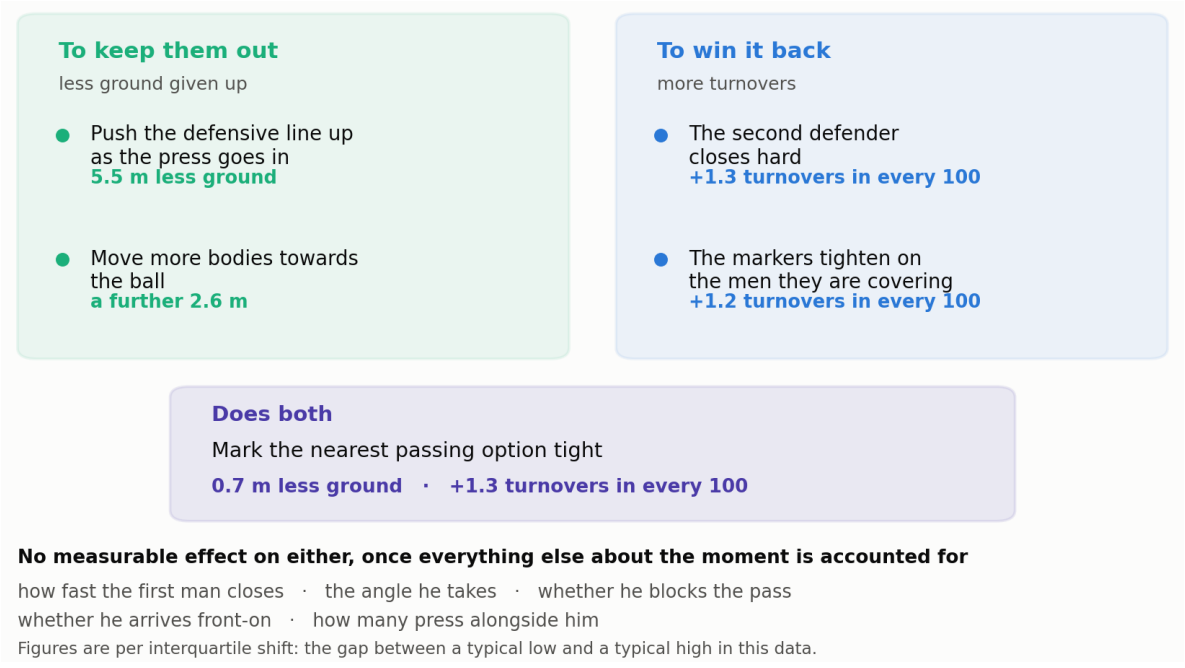


Figure 12. The paper on one page.

To keep an attack out, push the defensive line up as the press goes in, and move more bodies towards the ball rather than holding position. Together those are worth about eight meters of ground over five seconds. To win the ball back, the second defender in must close hard and the markers must tighten on the men they are covering. Those two actions win the ball, but they do not keep the attack out. Marking the nearest passing option tightly is the one instruction that serves both aims.

6.5 What this changes about measuring pressing

Three things, if the result holds up in other leagues and larger samples.

- Pressing metrics built from what individual defenders do near the ball are measuring the visible edge of something they are not measuring. They will correlate with outcomes, because the edge moves with the whole, but they will mislead anyone who treats them as levers.
- A single measure of press success is not enough. Ground and possession have different causes, and a metric that averages them will attribute a block's territorial success to the wrong players or score it as failure outright.
- And positional data alone is not sufficient for causal questions about pressing. The whole section 4.1 is a demonstration that an answer built on where players stand can invert once we add how they are moving. That is an argument for continuous tracking in this kind of work, and it is the clearest practical finding in the paper.

7 Future work

This study is small, and we would rather name what it cannot reach than imply otherwise. Each of the following is a piece of work we would like to do or to see done, and we would welcome collaboration on any of them.

Chance quality rather than territory. Our outcomes are ground and possession. Neither says whether the attack ended up dangerous. We built an on-ball-value outcome for exactly this and could not model it at the five-second resolution, as section 2.4 explains. Aggregating to the possession or the passage of play, rather than the individual press, is the obvious next attempt, and it is the one we would start with.

More matches, more leagues, more of the season. Thirty matches from six clubs in one league over part of one season is enough to detect effects of the size we report and not enough to say they generalize. Serie A in 2025/26 has its own pressing conventions. Whether the line's step is worth 5.5 m in a league that defends differently is an open question, and a straightforwardly answerable one with more data.

The attacking side of the same question. We measured what the defending team does. The mirror study which attacking movements cause a press to fail also uses the same design as the treatment vector moved to the other team, and section 6.3 sets out the specific hypothesis we would test first.

Learned representations of the situation. Our picture of the moment is handcrafted: a few dozen measurements were chosen because a coach would recognize them. A graph neural network over the tracking frames could capture configurations we did not think of.

Heterogeneity. Everything here is an average. Whether the line's step is worth more in one's own third than in the opponent's, or against one kind of attack rather than another, is a question of the same form and needs a larger sample to answer honestly.

Chain effects. We estimate the effect of a press on the five seconds that follow it. Whether a block that denies ground repeatedly accumulates over a half, an advantage is a different question at a different timescale, and arguably the one a manager would most like to answer.

8 Conclusion

We set out to build a causal framework for evaluating pressing efficiency on broadcast tracking data, and to apply it to the actions the game asks of the player nearest the ball.

Detecting presses from movement rather than from annotation gives 15,203 presses in 30 Serie A matches. Holding the situation fixed with double machine learning, using both where the players are and how they are moving, none of the five actions that player controls changes what happens next. The run direction that looks like the largest effect on the pitch when the situation is a snapshot disappears entirely once tracking data vector is in the picture.

What does change the outcome sits behind the ball and divides in two. The defensive line stepping up and defenders converging on the ball prevent ground and it is worth about eight meters between them. The second defender closing hard and the markers tightening win the ball back. Marking the nearest passing option tightly is the only action that helps on both.

None of this is an argument against pressing. It is an argument about where pressing happens, and it agrees with where the modern game has already moved: a press is something a team does, it is done mostly behind the ball, and it is two jobs rather than one.

References

Andrienko, G., Andrienko, N., Budziak, G., Dykes, J., Fuchs, G., von Landesberger, T. and Weber, H. (2017). Visual analysis of pressure in football. Data Mining and Knowledge Discovery 31(6), 1793-1839. doi:10.1007/s10618-017-0513-2

Anzer, G. and Bauer, P. (2022). Expected passes. Data Mining and Knowledge Discovery 36(1), 295-317. doi:10.1007/s10618-021-00810-3

Bauer, P. and Anzer, G. (2021). Data-driven detection of counterpressing in professional football. Data Mining and Knowledge Discovery 35(5), 2009-2049. doi:10.1007/s10618-021-00763-7

Bauer, P., Anzer, G. and Shaw, L. (2023). Putting team formations in association football into context. Journal of Sports Analytics 9(1), 39-59. doi:10.3233/jsa-220620

Cameron, A. C., Gelbach, J. B. and Miller, D. L. (2008). Bootstrap-based improvements for inference with clustered errors. Review of Economics and Statistics 90(3), 414-427. doi:10.1162/rest.90.3.414

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

Decroos, T., Bransen, L., Van Haaren, J. and Davis, J. (2019). Actions speak louder than goals: valuing player actions in soccer. Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 1851-1861. doi:10.1145/3292500.3330758

Fernandez-Navarro, J., Fradua, L., Zubillaga, A., Ford, P. R. and McRobert, A. P. (2016). Attacking and defensive styles of play in soccer: analysis of Spanish and English elite teams. Journal of Sports Sciences 34(24), 2195-2204. doi:10.1080/02640414.2016.1169309

Goes, F. R., Meerhoff, L. A., Bueno, M. J. O. et al. (2021). Unlocking the potential of big data to support tactical performance analysis in professional soccer: a systematic review. European Journal of Sport Science 21(4), 481-496. doi:10.1080/17461391.2020.1747552

Hirano, K. and Imbens, G. W. (2004). The propensity score with continuous treatments. In Applied Bayesian Modeling and Causal Inference from Incomplete-Data Perspectives, 73-84. doi:10.1002/0470090456.ch7

Imai, K. and van Dyk, D. A. (2004). Causal inference with general treatment regimes: generalizing the propensity score. Journal of the American Statistical Association 99(467), 854-866. doi:10.1198/016214504000001187

Link, D., Lang, S. and Seidenschwarz, P. (2016). Real time quantification of dangerousity in football using spatiotemporal tracking data. PLOS ONE 11(12), e0168768. doi:10.1371/journal.pone.0168768

MacKinnon, J. G. and Webb, M. D. (2018). The wild bootstrap for few (treated) clusters. The Econometrics Journal 21(2), 114-135. doi:10.1111/ectj.12107

Merhej, C., Beal, R., Matthews, T. and Ramchurn, S. (2021). What happened next? Using deep learning to value defensive actions in football event-data. Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, 3394-3403. doi:10.1145/3447548.3467090

Pappalardo, L., Cintia, P., Rossi, A. et al. (2019). A public data set of spatio-temporal match events in soccer competitions. Scientific Data 6(1). doi:10.1038/s41597-019-0247-7

Rein, R. and Memmert, D. (2016). Big data and tactical analysis in elite soccer: future challenges and opportunities for sports science. SpringerPlus 5(1). doi:10.1186/s40064-016-3108-2

Robinson, P. M. (1988). Root-N-consistent semiparametric regression. Econometrica 56(4), 931-954. doi:10.2307/1912705

Taki, T. and Hasegawa, J. (2000). Visualization of dominant region in team games and its application to teamwork analysis. Proceedings of Computer Graphics International 2000, 227-235. doi:10.1109/cgi.2000.852338

Appendix A The three pictures of a moment

The three pictures of section 5, in full. Each contains the one above it, so a number that changes between them changes because of what was added and nothing else.

Picture	What it contains
Snapshot	Fatigue and baselines for presser and carrier (cumulative distance, high-speed running in the last five minutes, top-speed baseline); ball x and y; distance presser to carrier; defenders within 15 m; carrier control time, local space, open passing lanes and orientation alignment; opponent hull area and attacker density; carrier press count in the last five minutes; game state; minute; players on the pitch for each team; play pattern. The counter press flag is excluded from all three, for reasons in section 5.

Picture	What it contains
+ movement	Ball speed and forward velocity; carrier speed and forward velocity; each team's mean forward velocity, excluding its own subject; rate of change of the defensive hull.
+ pitch context	Backline depth, ball to backline, defensive line spread, distance to goal, nearest teammate to carrier, goal-side attackers, attackers and defenders near the ball, the presser's own gap, distance to ball and goal-side indicator. Areas: defensive hull excluding the presser, and the overlap of the two hulls. Possession: seconds elapsed and events so far. Team identity for both sides.

Section 6 uses a different split again, described there: the attacking side in full, the first presser as a control, and the defending team's configuration one second before the trigger. It holds 39 numeric and 3 categorical variables.

Appendix B Every specification

All seventeen specifications reported in this paper, with the sample each was estimated on and the out-of-fold fit of its outcome model. Progression over five seconds except where the offset outcome is named.

Specification	What it adds	n	R2(Y|W)
S1a	The five first-defender actions, against a snapshot	15,203	0.180
S3	S1a plus ball speed and forward velocity in W	15,203	-
S4	A2 and the raw heading as co-treatments	15,203	-
S5	A2 replaced by heading, lane distance, approach angle	15,203	-
Snapshot	Section 5, picture one	15,203	0.180
+ movement	Section 5, picture two	15,202	0.264
+ pitch context	Section 5, picture three	15,202	0.289
PRE	Team actions, second before the trigger	14,350	0.289

Specification	What it adds	n	R2(Y|W)
PRE+PRIOR	Pre-trigger, with the preceding second's ball movement in W	13,930	0.285
POST	Team actions, second after the trigger	12,262	0.212
POST+PRIOR	Post-trigger, with the preceding second in W	11,632	0.210
POST+CONC	Post-trigger, with the same second in W (a lower bound)	11,632	0.232
PRE, ball won	Team actions before the trigger, against possession	14,350	0.105
POST, ball won	Team actions after the trigger, against possession	12,262	0.065
Combined x3	The single three-level scale at 5, 10 and 15 m	14,350	0.125-0.147

Appendix C Pitch conventions and detector settings

Every geometric quantity in the paper is computed after the coordinate conventions below are applied. Getting one of these wrong inverts a directional result silently, which is why each is checked by a known-answer test rather than by inspection.

Convention	Value
Pitch dimensions	105 x 68 m
Origin	Centre spot; x from -52.5 to +52.5, y from -34 to +34
Units	Meters and meters per second, never normalized
Sampling	25 Hz; one frame every 0.04 s
Attacking direction	Flipped per press so the carrier always attacks toward low x
Thirds	Defensive, middle and attacking at x = -17.5 and +17.5, from the pressing team's point of view
Smoothing	Savitzky-Golay, 11 frames, order 2, applied before velocities are differenced

Convention	Value
Ball coverage	63% of all frames; 92.9% of on-ball event frames

Parameter	Value	Role
Closure threshold	2.0 m/s	Minimum closing rate on the carrier for a press to be recorded
Onset distance	15.0 m	Maximum distance from the carrier at onset
Sustain	0.4 s	How long both conditions must hold
Re-arm	3.0 s	Minimum gap before the same player can trigger again
Dedup window	1.5 s	Collapses repeated detections of one approach
Zone gate	4.5 to 8.0 m	Arrival gate, by third of the pitch
Maximum duration	10.0 s	Beyond this the episode is not a press
Outcome horizon	5.0 s	Primary window, closed at the turnover
Sampling rate	25 Hz	Tracking frames per second
Detected	25,830	Presses found across the 30 matches
Event feed	10,451	Hand-coded Pressure events in the same matches
Smoothing	11 frames, order 2	Savitzky-Golay, applied before velocities are differenced

The closing threshold was settled by sweeping it rather than by convention. The same three conditions define a supporting presser in B2, which is what makes that measurement comparable with A5.

Appendix D Full model output

Every coefficient reported in the paper, with its interval, its p-value and the variation left in the action after the situation is accounted for. Effects are per interquartile shift of the action. Ground is in meters of progression allowed; ball won is in percentage points. An action with variation left below 0.10 would be reported as unanswerable; none in these tables falls below it.

D.1 The player nearest the ball, against three pictures of the moment

Picture	Effect (m)	95% interval	p	Variation left
A snapshot	-5.25	-5.93 to -4.57	0.001	0.56
+ movement	-1.11	-1.82 to -0.40	0.006	0.38
+ pitch context	-0.36	-1.06 to +0.34	0.311	0.38

D.2 The rest of the team, against ground conceded

Action	Effect (m)	95% interval	p	Variation left
B5 line pushes up	-5.47	-6.34 to -4.61	0.001	0.27
B3 defenders converging	-2.65	-3.17 to -2.14	0.001	0.55
B4 shape tightens	-0.69	-1.19 to -0.20	0.011	0.59
presser's run direction	-0.35	-1.23 to +0.54	0.461	0.15
B1 second defender closes	-0.16	-0.83 to +0.50	0.630	0.68
B7 markers tighten	-0.11	-0.46 to +0.24	0.526	0.96
B6 options shadowed	+0.20	-0.39 to +0.79	0.512	0.78
A1 presser's closing speed	+0.30	-0.28 to +0.88	0.324	0.51
B2 how many engage	+0.46	-0.12 to +1.05	0.118	0.69
B2 a second presser engages	+0.64	-0.46 to +1.74	0.246	0.76
A3 marking distance	+0.71	+0.27 to +1.15	0.003	0.69

D.3 The rest of the team, against the ball being won

Action	Effect (pp)	95% interval	p	Variation left
presser's run direction	+2.45	-0.32 to +5.23	0.084	0.15

Action	Effect (pp)	95% interval	p	Variation left
B1 second defender closes	+1.30	+0.17 to +2.43	0.025	0.68
B2 a second presser engages	+1.09	-1.42 to +3.60	0.380	0.76
B3 defenders converging	+0.83	-0.62 to +2.28	0.276	0.55
B6 options shadowed	+0.52	-0.96 to +1.99	0.484	0.78
A1 presser's closing speed	+0.41	-0.99 to +1.82	0.540	0.51
B2 how many engage	+0.33	-0.98 to +1.63	0.607	0.69
B7 markers tighten	-1.18	-1.99 to -0.36	0.011	0.96
A3 marking distance	-1.34	-2.61 to -0.07	0.043	0.69
B4 shape tightens	-1.43	-2.65 to -0.22	0.013	0.59
B5 line pushes up	-1.47	-3.13 to +0.19	0.080	0.27

Appendix E Reproducibility

Every figure in this paper is drawn from a stored result table rather than recomputed at draw time, so a number in the text and the same number in a figure cannot drift apart. A verification script restates each figure quoted in the text as an assertion against the table it came from, covering every coefficient, interval, p-value, overlap figure, sample size and placebo ratio reported here, and it is run before each revision of the manuscript.

Fold assignments are fixed by a stored seed and travel with the panel, so every arm of the study splits the data identically. Bootstrap replications are 999 throughout with a fixed seed. The feature layer enforces its temporal wall with read-only arrays views. The estimation code, the feature layer and the unit tests are available on request, subject to the data license described below.

Data and acknowledgements

This research was conducted with data provided by Hudl as part of the Hudl Performance Insights Research Competition. The data remains the property of Hudl and is not redistributed with this work.