

Entry Zones in Football

Estefania Vidal¹

The study of entry zones corresponds to the characterisation of the way in which teams enter the last part of the field. Since most shots are taken in the last portion of the field, the percentage of possession that produce shots increments with how far a team penetrates in their opponents area. Figure 1 shows how the percentage of possessions producing shots increases with the field advance. From there it can be seen that a possession that reaches the last third of the field has doubled its probability of producing shots in comparison to an average possession.

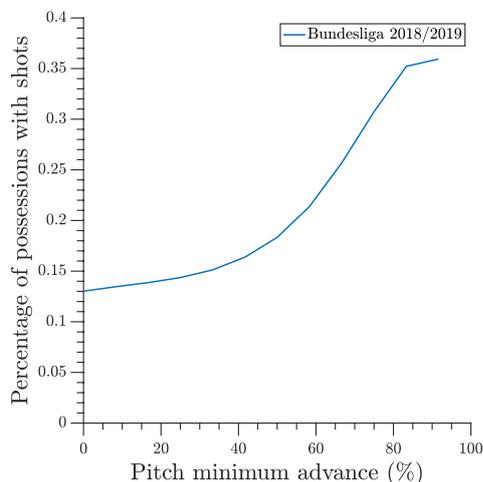


FIGURE 1: PERCENTAGE OF POSSESSIONS WITH SHOTS GIVEN THE MINIMUM FIELD ADVANCE OF THE POSSESSION

The data for this study was collected and provided by StatsBomb Services Ltd. from the season 2017-18 of 5 European leagues (La Liga, Serie A, Premier League, Bundesliga, and Ligue 1). The entire dataset amounted to 340847 possessions, as identified by StatsBomb. Of this dataset 89379 possessions entered the last quarter of the field, these possessions were identified as having at least one event starting in the last quarter following an event that did not occur in the last quarter. This ruled out set pieces and possession that started and never left the last quarter. All further analysis was performed

on this reduced dataset, which will be referred as possession entering the offensive zone.

From the possessions entering the offensive zone, 29% produced at least one shot, and from these possessions producing shots, 11% ended in goals.

From the different ways in which a team can enter the offensive zone, this analysis focused on the events identified in the dataset as “Carry” and as “Pass”. If during a possession the offensive zone was entered more than once, the type of the first entry was considered.

	Pass	Carry	Total
All events	64839	24540	89379
Shots	16697	8894	25591
Goals	1872	1056	2928
Shot %	25,8	36,2	28,6

FIGURE 2: TYPE OF ENTRY TO THE OFFENSIVE ZONE FOR ALL POSSESSIONS, POSSESSIONS PRODUCING AT LEAST ONE SHOT. AND POSSESSIONS ENDING IN GOALS

The initial division between type of zone entries is shown in Figure 2. From all possessions entering the offensive zone, carries proved to be more effective, producing shots 36% of the time, versus a 26% of the time for pass entries. Nevertheless, pass entries are much more common, amounting to 73% of all the offensive zone entries.

Carry entries

Analysis of the starting and ending locations of carry entries to the offensive zone was performed using gaussian kernel density estimation, to detect tendencies in the events. Figure 3 shows the probability density function of the starting and ending locations of the entry carry.

¹ @tefaidal

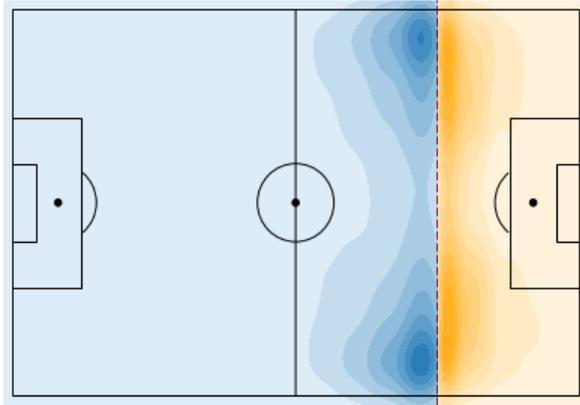


FIGURE 3: STARTING (BLUE) AND ENDING (ORANGE) LOCATION OF ALL CARRIES THAT ENTER THE OFFENSIVE ZONE

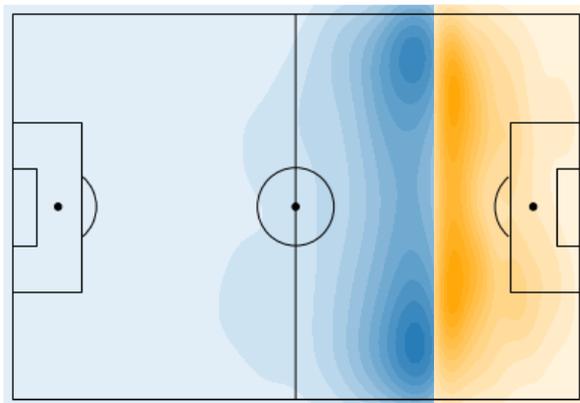


FIGURE 4: STARTING (BLUE) AND ENDING (ORANGE) LOCATION OF CARRIES THAT ENTER THE OFFENSIVE ZONE AND END IN GOALS

A marked tendency towards the sidelines can be observed with more than 40% of the events occurring in the last 10 meters to the side, and only 10% in the 20 meters closer to the centre. The events are characterised by being short (58% advance less than 20 meters in the horizontal direction) and moving mostly straight forwards (40% under 15° inclination as measured with respect to field's long axis).

Figure 4 shows the same analysis applied to the entry carries that produce goals. Even though this sample also shows a tendency towards the sidelines, the tendency is less strong and the peaks (most likely to happen locations) occur closer to the centre compared to the all events sample. The ending locations also show a shift in distribution, with the final location occurring further down the pitch and closer to the centre. For the possessions ending in goals, the length of the entry carry also shows a shift towards longer carries, with 50% of carries

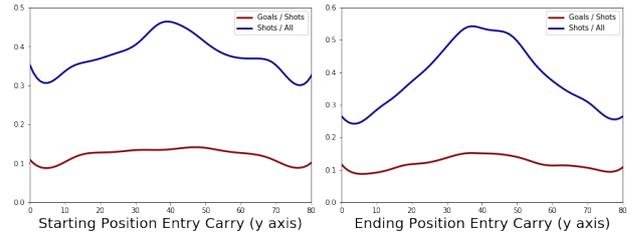


FIGURE 5: CONDITIONAL PROBABILITY OF A POSSESSION PRODUCING A SHOT (BLUE LINE) AND OF A SHOT BEING A GOAL (RED LINE) GIVEN THE STARTING POSITION (LEFT) AND ENDING POSITION (RIGHT) OF THE ENTRY CARRY

advancing the pitch between 10 and 30 meters.

These probability density functions can be used to calculate the conditional probability of a possession producing shots given the starting or ending position of the entry carry. This data is presented in Figure 5 and shows an increasing probability of a possession producing shots the closer it finishes to the centre of the pitch. As high as half of the possessions entering the offensive zone with carries ending in the middle 20 meters produce shots.

Even though the shot probability shows a high dependence on the location of the entry to the

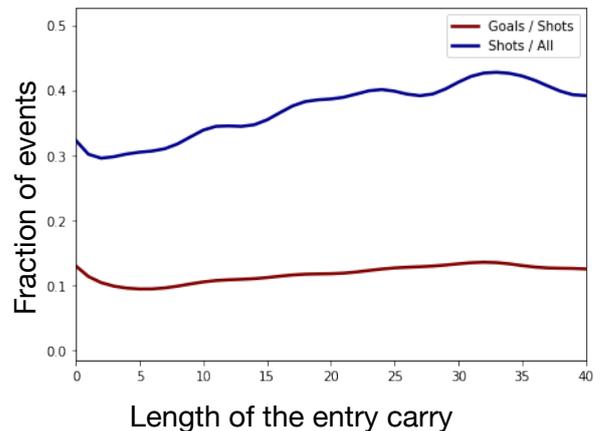


FIGURE 6: FRACTION OF POSSESSIONS PRODUCING SHOTS (BLUE) AND OF SHOTS PRODUCING GOALS (RED) GIVEN THE LENGTH OF THE ENTRY CARRY

offensive zone, the probability of scoring a goal (given that a shot was taken) remained practically constant with the exception with entries too close to the sidelines, where the probability of goal scoring decreased.

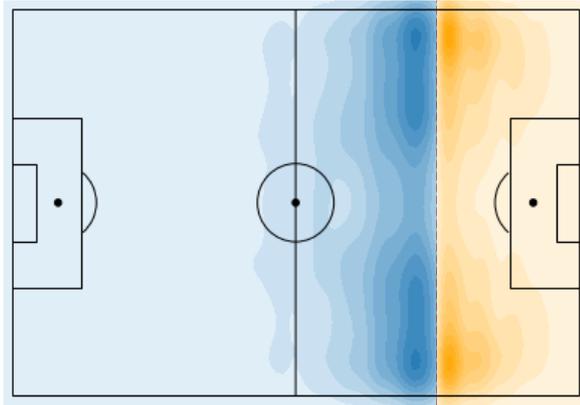


FIGURE 7: STARTING (BLUE) AND ENDING (ORANGE) LOCATION OF ALL PASSES THAT ENTER THE OFFENSIVE ZONE

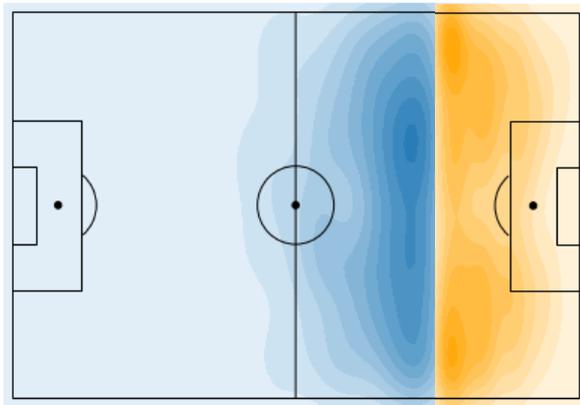


FIGURE 8: STARTING (BLUE) AND ENDING (ORANGE) LOCATION OF PASSES THAT ENTER THE OFFENSIVE ZONE AND END IN GOALS.

The length of the entry carry also showed correlation with the shot probability (see Figure 6), possessions entering the offensive zone with carries longer than 15 m have a higher probability of producing shots.

Pass Entries

The probability distribution function for the starting and ending locations for the passes that enter the offensive zone are presented in Figure 7. Unlike carry entries, the starting location of entry passes are distributed more evenly across the side of the pitch, while the ending location of passes is more sharply distributed towards the sidelines. This shows that passes tend to open the field, being played towards the sidelines, in contrast with carries, which are played from the sidelines towards the centre.

Filtering the entry passes to those that end in goals (see Figure 8) shows a shift in the starting location distribution towards the

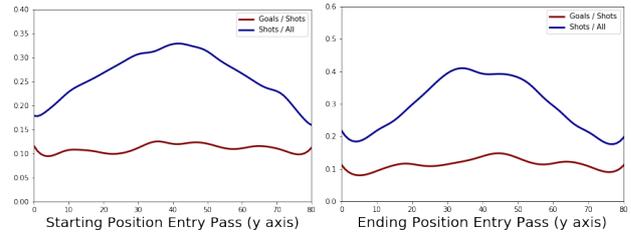


FIGURE 9: CONDITIONAL PROBABILITY OF A POSSESSION PRODUCING A SHOT (BLUE LINE) AND OF A SHOT BEING A GOAL (RED LINE) GIVEN THE STARTING POSITION (LEFT) AND ENDING POSITION (RIGHT) OF THE ENTRY PASS

centre. The ending location distribution shows that shorter passes are more commonly towards the sidelines, while longer passes reach further down the field.

The conditional probability of possessions ending in shot and of shots ending in goals given the position of the entry pass is presented in Figure 9. How dangerous a possession is increases with how close it is to the middle of the field. In particular for passes, only those starting in the middle 20 meters of the field or entering the offensive zone in the middle third of the pitch present an above average (higher than 30%) shot probability.

The length distribution for passes entering the offensive zone did not present great variability between all possessions and possessions ending in goals. Most passes advance the field in the horizontal axis between 10 and 30 meters, independent of their risk. Therefore, the fraction of possessions producing shots

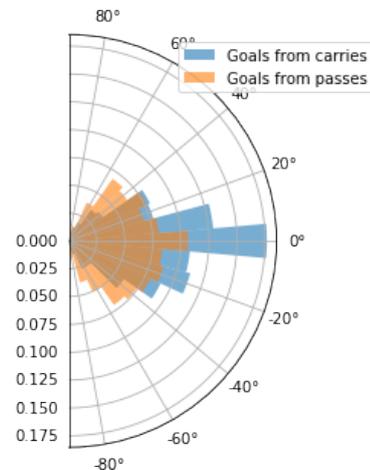


FIGURE 10: ANGLE DISTRIBUTION WITH RESPECT TO THE LONG FIELD AXIS, OF POSSESSIONS ENDING IN GOALS, DIVIDED BY THEIR TYPE OF ENTRY TO THE OFFENSIVE ZONE. POSSESSIONS ENTERING THROUGH CARRIES IN BLUE, AND THROUGH PASSES IN ORANGE.

remained independent of pass length for passes longer than 17 meters.

The distribution of the entry angle measured with respect to the long field axis for passes and carries entering the offensive zone and ending in goals are presented in Figure 10. This highlights one of the major differences between both types of entries. While carries direction is mostly straight forward towards the end of the pitch, passes have a much wider and homogenous angle distribution. That is why passes go from a homogenous distribution across the pitch, to a mark sideline tendency after entering the offensive zone. While carries maintain their sideline tendency (only slightly more homogenous towards the centre) after entering the offensive zone. A marked difference in entry angle between all possessions and possessions ending in goals was not observed.

Re-entries

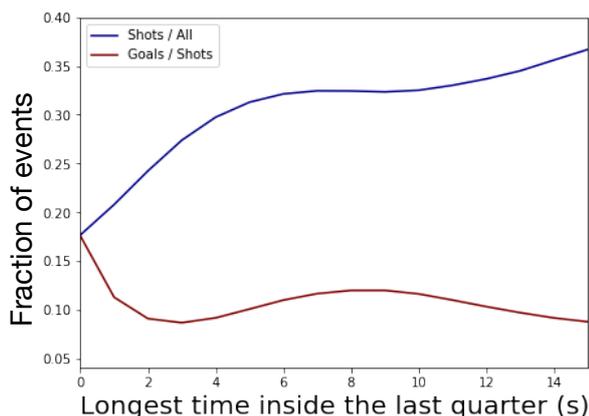


FIGURE 11: FRACTION OF POSSESSIONS PRODUCING SHOTS (BLUE) AND FRACTION OF SHOTS PRODUCING GOALS (RED), GIVEN HOW LONG THE POSSESSION WAS IN THE OFFENSIVE ZONE

How effective a possession is depends heavily on how long it remains inside the offensive zone. The longer a team remains in the offensive zone, more likely it is that they would shoot, nevertheless with more time attacking the probabilities of these shots to end in goals decreases. This effect can be seen in the conditional probabilities presented in Figure 11. This can be understood in terms of teams rearranging and organising their defence. The longer the attacking team remains in the offensive zone, more anxious the players get and it is therefore more likely that they would try to shoot. On the other hand, this extra time gives the defence time to bring more players inside their defensive zone and to cover the

spaces in a more organising way, therefore decreasing how effective those shots are. Figure 11 shows how after 8 seconds of a possession inside the offensive zone, the probability of a shot ending in goal starts to decrease. Because of this, after 12 seconds in the offensive zone it is recommended to go out of the zone and reorganise the attacking efforts.

Build-ups

As a particular case of possessions entering the offensive zone, we now shift our attention to build-ups. This type of possession will be defined as a possession that starts in the first quarter of the pitch and that crosses the last quarter, i.e., a possession that starts in the defensive zone and enters the offensive zone. Under this definition, counterattacks are a special case of build-ups.

A careful analysis of the events leading up to entering the offensive zone shows no correlations between the amount or type of events, that is how many carries and passes, that were needed to reach the offensive zone and the probability of a possession ending in goals.

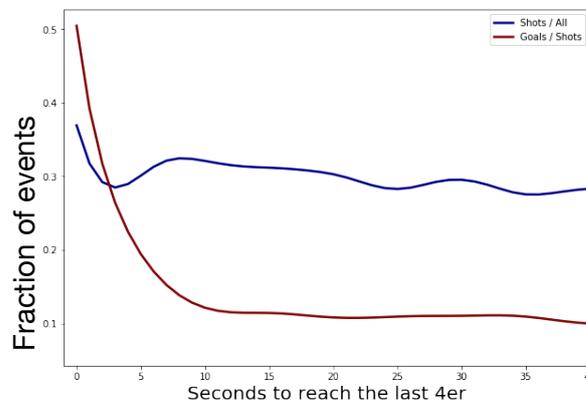


FIGURE 12: FRACTION OF BUILD-UPS PRODUCING SHOTS (BLUE) AND FRACTION OF SHOTS PRODUCING GOALS (RED), GIVEN HOW LONG IT TOOK TO REACH THE OFFENSIVE ZONE

However, the time it takes for a build-up to reach the offensive zone did show a correlation, as presented in Figure 12. Build-ups that reach the offensive zone quickly (less than 10 s) have a higher probability to score a goal if they produce shots. While the probability to produce shots decays slightly with time, it remains fairly constant. After ten second the goal probability remains the same and it is equivalent to the goal probability of a

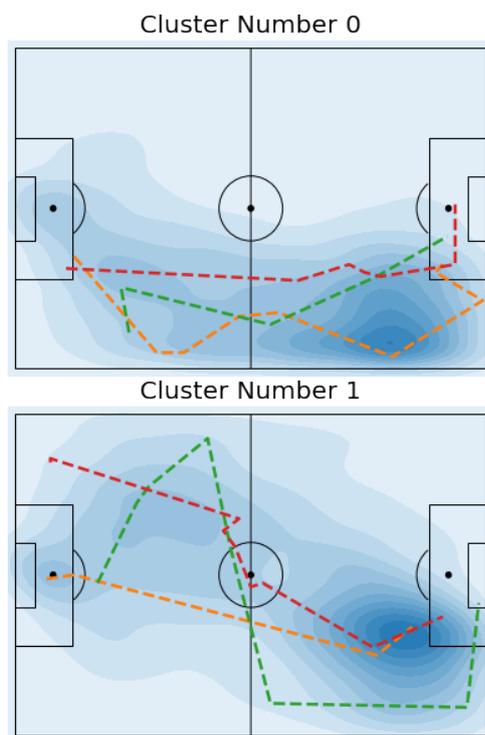


FIGURE 13: TWO TYPES OF BUILD-UPS ENTERING THE OFFENSIVE ZONE THROUGH THE RIGHT. BLUE COLOUR SHOWS THE MORE COMMON LOCATIONS FOR THE BALL PATHS, WHILE THE DOTTED LINES SHOW SOME PATH EXAMPLES.

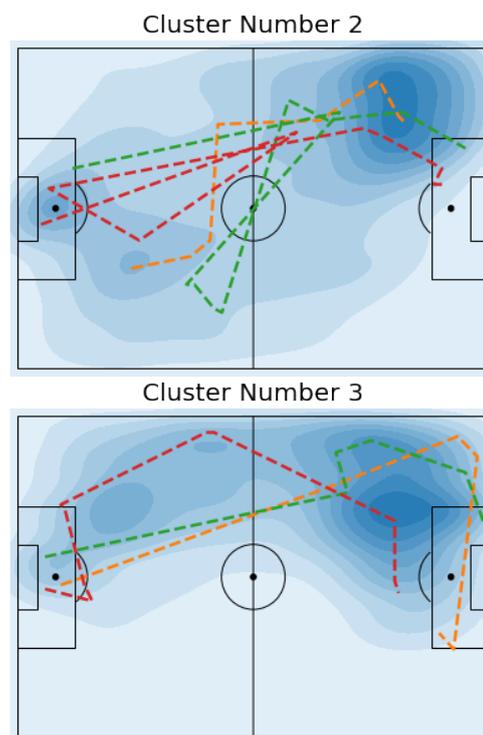


FIGURE 14: TWO TYPES OF BUILD-UPS ENTERING THE OFFENSIVE ZONE THROUGH THE LEFT. BLUE COLOUR SHOWS THE MORE COMMON LOCATIONS FOR THE BALL PATHS, WHILE THE DOTTED LINES SHOW SOME PATH EXAMPLES.

normal possession (not build-up) that reaches the offensive zone. This can be interpreted as losing the extra edge given by the build-up if it takes too long to reach the offensive zone, since the defenders would have enough time to also reach their positions and arrange their defence.

We used clustering analysis to identify the different types of build-ups. Clustering was performed using k-means++ and the distance between trajectories was calculated using Frechet distance² as implemented for Python by Jekel et al.³

The automatised process detected four main types of build-up. Two that enter the offensive zone mainly through the right side (Figure 13) and two that enter mainly through the left side (Figure 14). The first cluster (Cluster 0) is characterised by going quickly towards the right side and remaining close to the sideline

as the possession advances. The play moves closer to the centre of the pitch far into the offensive zone.

Cluster 1 moves closer to the centre in the first half, even with a slight shift towards the left side of the attacking area, to then enter the offensive zone towards the right, very close to the penalty box corner. The build-ups whose advance was mostly through the middle of the the pitch are in this cluster.

Cluster 2 is very symmetric to Cluster 0, with the difference that carries through the left tend to go more towards the sideline and reach the offensive zone heavily towards the left.

The final cluster corresponds to build-up through the left that advance mostly close to the sideline, but that move towards the centre in the last quarter and reach the offensive zone close to the penalty box. These

² M Maurice Fréchet. Sur quelques points du calcul fonctionnel. Rendiconti del Circol Matematico di Palermo (1884-1940), 22(1):1-72, 1906.

³ Jekel, C. F., Venter, G., Venter, M. P., Stander, N., & Haftka, R. T. (2018). Similarity measures for identifying material parameters from hysteresis loops using inverse analysis. International Journal of Material Forming.

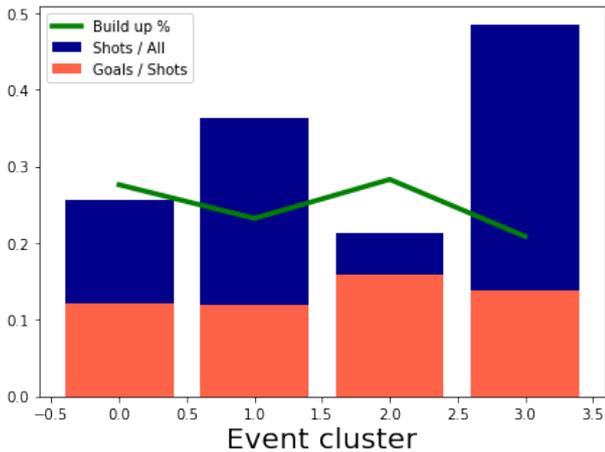


FIGURE 15: FRACTION OF BUILD UPS CLASSIFIED IN EACH CLUSTER (GREEN). FRACTION OF BUILD-UPS THAT PRODUCE SHOTS, GIVEN THEIR CLUSTER CLASSIFICATION IN BLUE, AND FRACTION OF SHOTS THAT END IN GOALS, GIVEN THEIR CLUSTER CLASSIFICATION IN ORANGE.

possessions tend to cross towards the middle of the pitch for better shot locations.

Figure 15 shows the fractions of build-ups classified in each cluster. The distribution was fairly uniform, but the two clusters that go more extremely towards the sidelines (Clusters 0 and 2) had the higher percentage of occurrence. Of all types of build-ups, types 1 and 3 had an above average probability of producing shots. These build-ups are more centred and enter the offensive zone close to the corner of the box. In particular, build-ups belonging to cluster 3, even though they enter through the left, they manage to successfully cross the field and shoot closer to the centre, thus destabilising the defence .

To finish up build-ups we now take a look at the characteristics of the last event before a shot. How much the last event advances the pitch is presented in Figure 16. It can be seen that the pass distribution presents two local maxima: at short backwards passes, and at long (between 10 and 20 m) forward passes. Carries on the other hand are much shorter and less often going backwards from the end of the pitch, compared to passes.

The left side of Figure 16 shows that 42% of last passes were back passes while only 21% of the last carry events moved backwards on the pitch. In a similar manner as the previous description, passes tend to be longer than shots, with 22% of the last passes being longer than 10 m, while only 9% of carries reached that length.

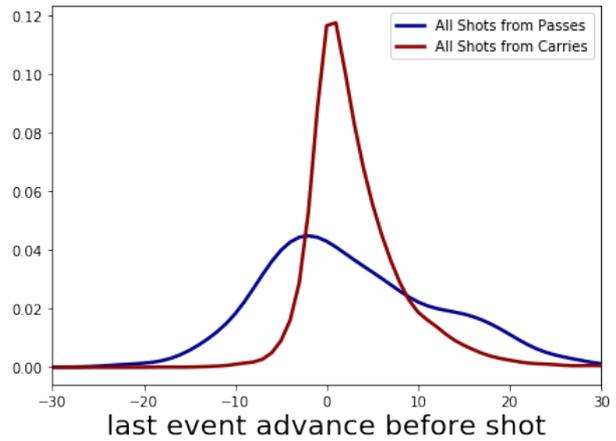


FIGURE 16: VERTICAL ADVANCE ALONG THE FIELD OF THE LAST EVENT BEFORE A SHOT. PASSES IN BLUE AND CARRIES IN RED. ALL TYPES OF BUILD-UPS.

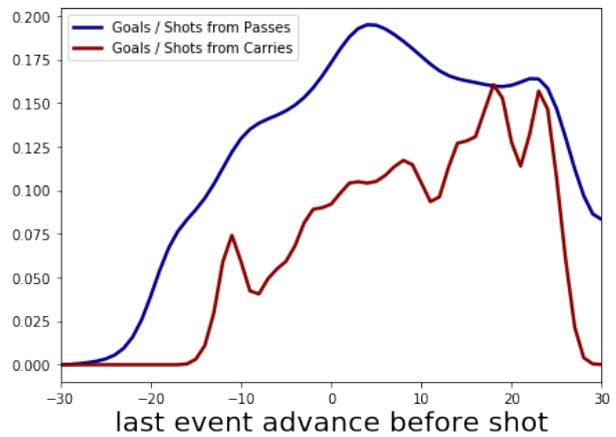


FIGURE 17: PROBABILITY OF A SHOT ENDING IN GOAL GIVEN THE VERTICAL ADVANCE OF THE EVENT BEFORE. PASSES IN BLUE AND CARRIES IN RED. ALL TYPES OF BUILD-UPS.

Finally, from the shots produced ending build-ups those with higher probability of ending in goals come from short forward passes (see Figure 17). If moving backwards is necessary for better shot location, then passing backwards is generally more effective than carry from the end-line. While for longer forwards distances passes and carries are equally effective.

Final remarks

Independent of carries or passes, playing to the middle continues to be the most effective way of producing shots. The more centred the entry event, the better.

The effects of sustained possession in the offensive zone decline fast in time. Because of

this, a hard reset, going out of the offensive zone, is recommended after ~12 seconds.

For build-ups, they are more effective when they are fast, taking less than 10 second to reach the last quarter. They should avoid long balls and the sidelines.

The explanation of some of these statistics is varied and cannot be easily separated without taking a further look into player positions. For example, we have shown that entering the offensive zone through carries is more effective than through passes, but the questions remains, is it more effective to carry the ball, because it drags defenders and disorganises the defence? or the players choose to carry the ball because they already observe that the field is open enough to allow such an entrance? Technology like player tracking and metrics to asses the openness of the pitch could help answer these questions.