Evaluating Playmakers in 9 Leagues:
Opportunity-adjusted chance creation

How to judge a playmaker?

We build a data-driven *model of the many* which estimates the number of chances created (CCs) that we should expect of the typical player, based on the opportunities given him, such as minutes on the pitch, crosses he gets to make, opportunities given him for passing in the final third, etc. This creates a sliding scale against which to benchmark all players as playmakers. Anyone who makes better use of these opportunities (by making more incisive passes, crosses) will outperform the typical by making more CCs than expected. This excess in opportunity-adjusted chance creation is our measure of playmaking talent, and the very best playmakers will register as outliers in excess CC.

We use public domain data from Opta Analyst to replicate the analyses over 8 different European leagues, 2025-26, plus a single legacy dataset from the previous season for one of the leagues. We build the *model of the many* using robust regression, which down-weights the few observations that do not follow trend. This gives a cleaner *model of the many* than OLS regression. But the main benefit is that robust regression clearly identifies outliers that do not follow trend. Usually, outliers are simply rejected from analyses as aberrant data due to coding error, unidentified exceptions of equipment failure, and so on. But in judging playmakers, (positive) outliers are to be valued as a central focus of interest.

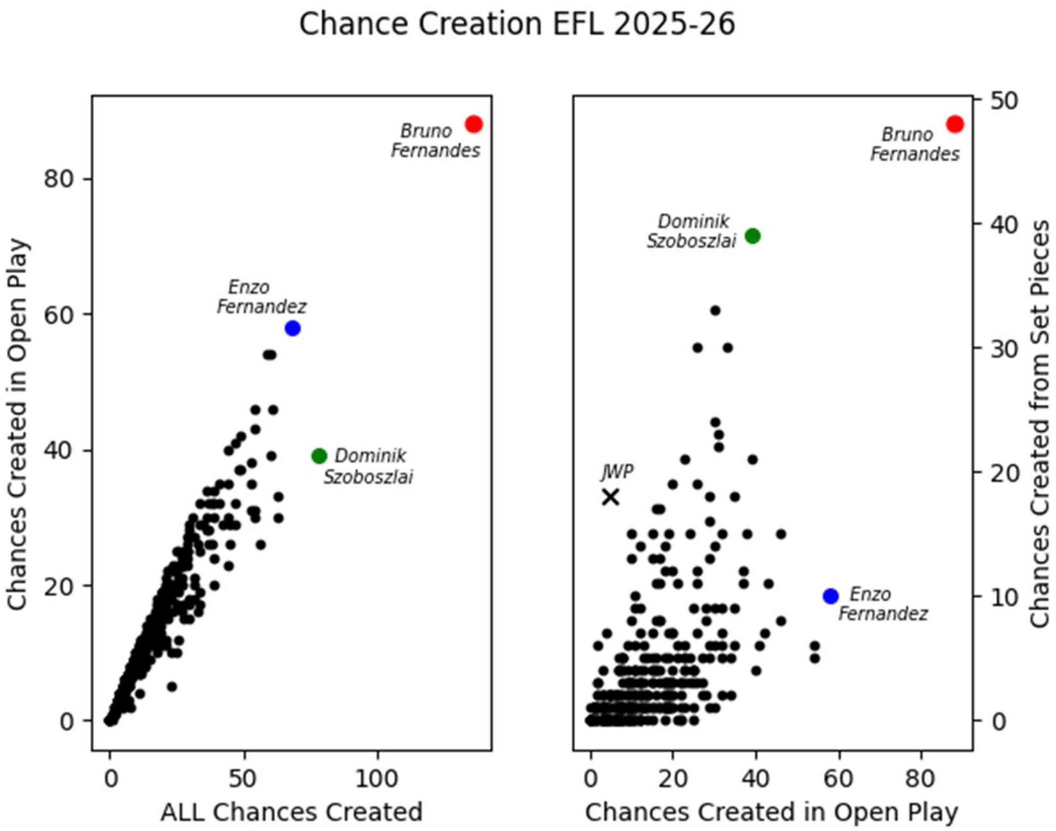
Bruno Fernandes' football boosts

One of the subplots of this year's EFL was Bruno Fernandes' 21 assists for the season, breaking the previous bests of de Bruyne (20), Henry (20), and Özil (19). To achieve such a feat you have to be an outstanding playmaker, but fortune has to be on your side too, because BF's expected assists for the season are recorded as only 12.32, suggesting that luck, excellent finishing, and / or indifferent goalkeeping helped give his numbers an *after-pass* boost. Then there has to be *opportunity*: BF played 90% of the season's minutes, which is high for players in forward positions; and he took most free kicks and 90 corners. His xA reduces to just 9.12 in open play when set pieces are subtracted. Finally, team-mates needed to help by putting BF in positions where he could create those chances, which Manchester United certainly did. For instance, BF made 871 passes in the final third, which was 10% more than the next highest in the league (Dominik Szoboszlai).

In view of these boosts, how remarkable was BF's underlying performance when adjusted for opportunity? And which other playmakers are worthy of note (who may not have been given the same opportunities)?

Chance Creation in open play

Set pieces are undoubtedly important. Arsenal won this year's EPL on the back of them, and being a dead ball specialist, such as James Ward-Prowse (marked in the right scattergram with an "X"), can certainly improve a CC tally: 5 in open play, augmented by 18 from set pieces. But most players don't get a chance to take free kicks or corners, which obviously skews the statistics in favour of those who do. Set pieces also require a different skill set to open play creation. Consequently, we restrict our attention to CC in open play, and leave set pieces for another stage.



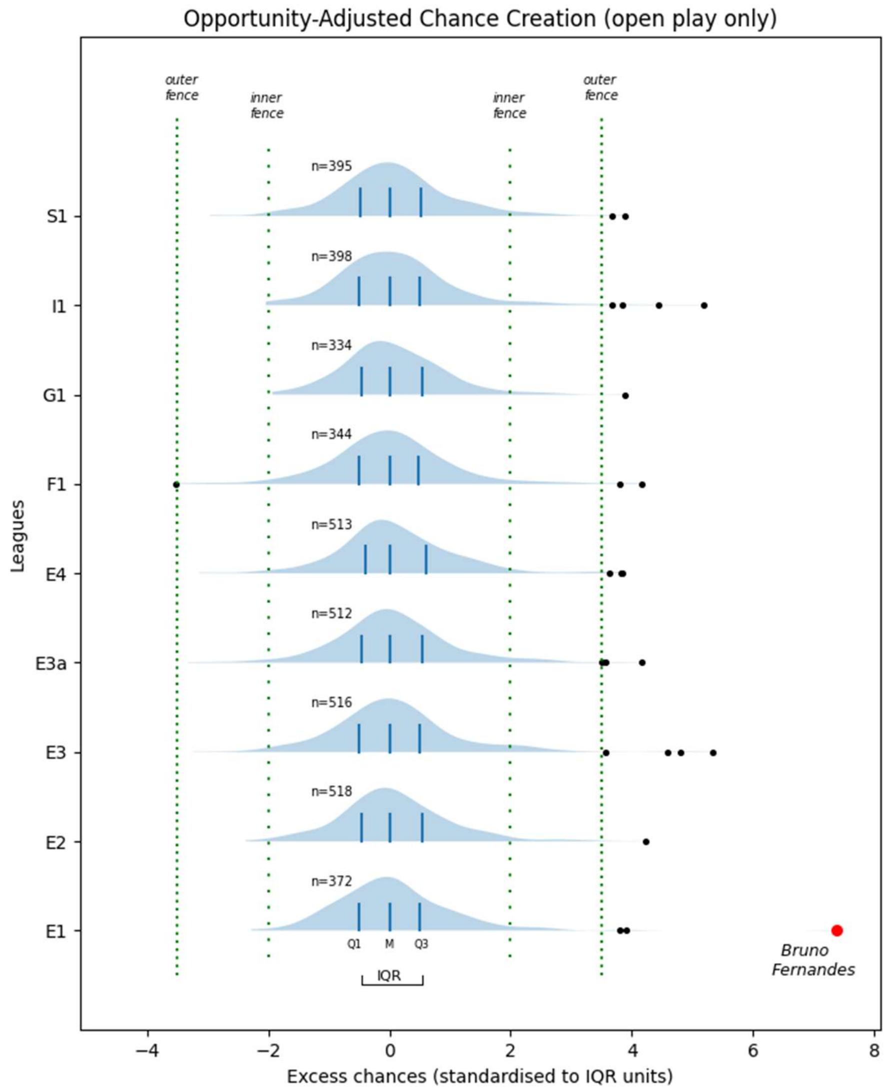
The left scatterplot did the rounds in social media and TV. But reworking the data as in the right scatterplot is more instructive. BF (●) had most CCs in open play. Next nearest was Enzo Fernandez (●), who had his own boost of even more minutes than BF. BF also gained more CCs from set pieces than anyone else – next nearest was Dominic Szoboszlai (●).

Model of the many

We build our *model of the many* using robust regression[1] for each league in turn to estimate the CC we should expect a player to typically achieve, given his number of minutes, number of passes in the final third, number of crosses, and number of passes *not* in the final third – all publicly available from Opta Analyst. All players played at least 450 minutes. The playmaker's job is to turn these *opportunities* into chances created. If a player does this better than average, he will have a positive regression residual (an excess of chances above expected). We use this residual to measure players' creativity in each of 9 leagues: The Big Five (E1, F1, G1, I1, S1 for EPL, Ligue 1, Bundesliga, Serie A, and La Liga) and the other English Football Leagues below E1 (= EPL). They are E2, E3, E4, plus E3 in the 2024-25 season, listed as E3a.

[1] Specifically, Tukey's biweight regression with c = 4.685, and hc3 correction for heteroskedasticty.

Distribution of players in each league, with outliers shown



This graph shows the smoothed distributions for all 9 leagues. What Tukey called **far outliers** are marked as filled black circles; they are defined as lying greater than 3 x IQR above Q3 (≈ 3.5 x IQR above the median) – similarly for far outliers below the median. These are much more extreme than the usual outliers found in Tukey boxplots, defined as lying 1.5 x IRQ above (below) Q3 (Q1). Because leagues have different numbers of games (max = 46, min = 34), the raw, expected, and excess CCs will naturally differ in magnitude. Measuring *Excess CC* in league-specific IQRs units, as in the graph, allows us to compare across leagues. Evidently Bruno Fernandes (in red) is the most extreme outlier in all leagues. Although his team certainly gives him more opportunities than most to create chances, once this boost is adjusted for, BF is still revealed to be exceptionally efficient at turning those opportunities into chances. In the next section are the top 10 players in each league, measured by Excess CC. Far outliers are highlighted in green.

Top ten playmakers in each league

English1	Mins	CC	ExpCC	Excess	Excess in IQRs
Bruno Fernandes	3066	88	47.02	40.98	7.40
Jérémy Doku	1784	54	32.34	21.66	3.90
Enzo Fernández	3121	58	36.92	21.08	3.80
Casemiro	2588	34	19.26	14.74	2.65
Bukayo Saka	2225	46	31.65	14.35	2.58
Florian Wirtz	2388	54	39.96	14.04	2.52
Cody Gakpo	2747	46	32.09	13.91	2.50
Yéremy Pino	2091	37	23.23	13.77	2.48
Marc Cucurella	2710	34	20.66	13.34	2.40
Callum Hudson-Odoi	1844	40	27.36	12.64	2.27

English2	Mins	CC	ExpCC	Excess	Excess in IQRs
Ephron Mason-Clark	2989	58	35.09	22.92	4.22
Jack Clarke	2380	48	29.37	18.63	3.44
Isaac Price	3642	52	33.76	18.24	3.36
Scott Twine	3321	50	32.83	17.17	3.17
Tommy Conway	3597	54	37.74	16.26	3.00
Edo Kayembe	2496	43	27.12	15.88	2.93
Tom Fellows	2430	41	25.82	15.18	2.80
Finn Azaz	3068	53	38.12	14.88	2.75
Femi Azeez	2810	49	34.79	14.21	2.63
Patrick Roberts	2208	42	28.23	13.77	2.55

English3	Mins	CC	ExpCC	Excess	Excess in IQRs
Rubin Colwill	2412	60	30.77	29.23	5.33
Patrick Kelly	2902	48	21.67	26.33	4.80
Mason Burstow	2518	44	18.87	25.13	4.59
Kamari Doyle	2596	39	19.40	19.60	3.57
Ben Osborn	2942	52	35.85	16.15	2.95
Joel Bagan	3631	57	41.20	15.80	2.88
Amario Cozier-Duberry	2651	43	27.78	15.22	2.77
Kyran Lofthouse	4014	54	39.15	14.85	2.71
Kyrell Lisbie	3212	37	22.77	14.23	2.59
Reeco Hackett	2895	38	23.96	14.04	2.56

English3a	Mins	CC	ExpCC	Excess	Excess in IQRs
Ben Wiles	3423	53	29.59	23.41	4.16
Ryan Barnett	2867	51	30.91	20.09	3.57
Jeremy Kelly	2793	40	20.26	19.74	3.51
Lee Evans	3523	46	29.91	16.09	2.86
Albie Morgan	3164	43	27.61	15.39	2.74
Joel Randall	2307	32	16.85	15.15	2.70
Jamie Donley	2762	44	29.27	14.73	2.62
John McAtee	2795	38	23.36	14.64	2.60
Callum Marshall	2677	35	20.44	14.56	2.59
Kamari Doyle	2128	35	21.07	13.93	2.48

English4	Mins	CC	ExpCC	Excess	Excess in IQRs
Mark Helm	3389	50	28.67	21.33	3.85
Bobby Kamwa	2716	45	23.90	21.10	3.81
Daniel Udoh	2841	41	20.89	20.12	3.64
Isaac Hutchinson	3156	40	21.24	18.76	3.39
Dion Pereira	1774	34	15.56	18.44	3.34
Ismeal Kabia	3021	42	23.85	18.15	3.29
Charlie Whitaker	2939	42	24.05	17.95	3.25
Jake Bickerstaff	2962	33	17.52	15.48	2.81
Nick Tsaroulla	3006	42	27.46	14.54	2.64
Jorge Grant	3230	40	27.29	12.71	2.32

French1	Mins	CC	ExpCC	Excess	Excess in IQRs
Matthieu Udol	3003	48	28.46	19.54	4.16
Yann Gboho	2630	53	35.14	17.86	3.80
Ilan Kebbal	2229	47	31.81	15.19	3.23
Giorgi Tsitaishvili	2457	36	21.97	14.03	2.98
Samir El Mourabet	2135	25	11.72	13.28	2.82
Adrien Thomasson	2690	41	27.97	13.03	2.76
Senny Mayulu	1681	32	20.85	11.15	2.36
Jean-Philippe Krasso	1090	21	10.03	10.97	2.32
Ludovic Blas	1541	31	20.12	10.88	2.30
Jonathan Clauss	2531	38	27.44	10.56	2.23

German1	Mins	CC	ExpCC	Excess	Excess in IQRs
Romano Schmid	2988	69	47.54	21.46	3.89
Jean-Luc Dompé	1136	30	13.95	16.05	2.91
Jean-Mattéo Bahoya	1329	27	12.70	14.30	2.59
Albert Sambi Lokonga	1712	25	10.76	14.24	2.58
Christian Eriksen	2419	38	25.11	12.89	2.34
Ridle Baku	2509	34	21.44	12.56	2.28
Yan Diomande	2476	47	34.83	12.17	2.21
David Raum	2556	39	27.91	11.09	2.01
Luis Díaz	2450	50	39.22	10.78	1.96
Arijon Ibrahimovic	2181	35	24.94	10.06	1.83

Italian1	Mins	CC	ExpCC	Excess	Excess in IQRs
Martin Baturina	1569	47	19.84	27.16	5.19
Nicolò Fagioli	2564	48	24.76	23.24	4.44
Kenan Yildiz	2837	64	43.92	20.08	3.84
Charles De Ketelaere	2188	58	38.74	19.26	3.68
Nikola Vlasic	3053	49	32.42	16.58	3.17
Mario Pasalic	1932	36	19.58	16.42	3.14
Juan Miranda	2573	43	28.71	14.29	2.73
Youssouf Fofana	2217	33	19.49	13.51	2.58
Sebastiano Esposito	2700	43	29.69	13.31	2.55
Andrea Cambiaso	2794	46	32.86	13.14	2.51

Spanish1	Mins	CC	ExpCC	Excess	Excess in IQRs
Pablo Fornals	2865	59	39.72	19.28	3.88
Kylian Mbappé	2604	60	41.74	18.26	3.67
Pedri	2107	54	40.31	13.70	2.76
Ander Barrenetxea	1769	37	23.36	13.64	2.75
Vinícius Júnior	2825	62	48.93	13.07	2.63
Mikel Oyarzabal	2714	37	24.58	12.43	2.50
Álvaro García	2105	34	22.38	11.62	2.34
Pere Milla	1986	28	16.77	11.24	2.27
Lucas Boyé	1959	27	15.80	11.20	2.26
Antony	2459	48	36.95	11.05	2.23

Mins: minutes played in the season (filter >=450 mins).
CC: Chances created in open play over the season.
ExpCC: Expected number of chances created, according to model.
Excess: Regression residuals (CC minus ExpCC).
Excess in IQRs: Excess rescaled to be measured in units of IQR, relative to the median.
(IQR = interquartile range = Q3 minus Q1.)

Regression Models for each league

	Mins x1000	Pass 1/3rd x100	Crosses x100	Pass 2/3rd x100	Const x1
English1	3.601	5.259	4.500	-1.251	0.807
English2	5.719	6.348	4.187	-1.293	-0.060
English3	3.044	4.411	6.608	-1.058	0.389
English3a	3.628	4.250	5.337	-1.059	0.001
English4	3.078	4.606	6.424	-1.349	0.500
French1	4.059	6.446	1.288	-1.253	-0.093
German1	4.599	5.534	4.169	-1.304	-0.128
Italian1	4.026	5.532	5.483	-1.164	-0.059
Spanish1	2.521	6.259	4.083	-1.111	0.388
English1	6.60	12.76	3.44	-12.25	1.54
English2	12.24	19.74	4.30	-13.29	-0.15
English3	4.68	9.66	5.11	-10.13	0.86
English3a	8.02	12.58	4.53	-12.64	0.00
English4	6.33	12.94	7.44	-12.39	1.21
French1	5.92	13.96	0.72	-13.00	-0.20
German1	8.25	18.28	2.46	-13.80	-0.24
Italian1	6.73	17.24	3.65	-11.04	-0.12
Spanish1	3.96	12.63	3.32	-10.16	0.77

The top half of the table (green) are the regression coefficients, and the bottom half (blue) are the corresponding t-values. The coefficients have been scaled for easier viewing. They are to be read as follows: if a player competing in the EPL (English1) plays for 1000 minutes, makes 100 passes in the final third, makes 100 crosses, and makes 100 crosses NOT in the final third, then we expect him to create:

$$3.601 + 5.259 + 4.500 - 1.251 + 0.807 = 12.915 \text{ chances}$$

If instead he makes 200 passes in the final third, we expect him to create an additional 5.259 chances, and so on. Note that minutes *per se* do not have much bearing on chance creation, and that the more passes *not* in the final third, the *fewer* the chances we expect to be created. R^2 goodness of fit for the 9 leagues models were: 0.816, 0.859, 0.799, 0.785, 0.822, 0.859, 0.826, 0.824, 0.837, in turn.

A single model of Chance Creation?

The coefficients (highlighted in green) across the leagues are remarkably similar. To appreciate quite how similar, suppose we mistakenly applied F1's regression coefficients to E2 players, for instance. How much difference would it make to CC estimates for each player? The table below shows all the Pearson correlation coefficients of expected CCs when a 'wrong' set is applied, correlated with expected CCs using the correct set. Specifically, row 2 column 6 answers our question. Wrongly apply the F1 coefficients to E2 players and the resulting estimates have an extremely high correlation of 0.9903 with the correct estimates (E2 weight applied to E2 players).

The other boxed correlation in the table is the *vice versa*: wrongly applying E2 weight to F1 players. Unlike in a standard correlation matrix, these figures differ slightly.

Players (in the 9 leagues)	Model coefficients (from each of the 9 leagues)								
	E1	**E2**	**E3**	**E3a**	**E4**	**F1**	**G1**	**I1**	**S1**
E1	1	0.9881	0.9977	0.9984	0.9881	0.9910	0.9982	0.9968	0.9943
E2	0.9776	1	0.9712	0.9842	0.9246	0.9903	0.9918	0.9932	0.9793
E3	0.9976	0.9858	1	0.9986	0.9939	0.9834	0.9943	0.9956	0.9919
E3a	0.9986	0.9908	0.9985	1	0.9874	0.9880	0.9980	0.9980	0.9924
E4	0.9920	0.9736	0.9964	0.9930	1	0.9693	0.9857	0.9871	0.9830
F1	0.9906	0.9948	0.9824	0.9870	0.9574	1	0.9946	0.9947	0.9966
G1	0.9981	0.9945	0.9946	0.9979	0.9762	0.9940	1	0.9987	0.9932
I1	0.9958	0.9954	0.9937	0.9971	0.9694	0.9933	0.9987	1	0.9951
S1	0.9941	0.9899	0.9882	0.9892	0.9705	0.9961	0.9937	0.9953	1

The conclusion is that model coefficients are highly interchangeable between leagues, with the possible exception of F1 with E4 (highlighted in yellow). Highest coherence is between E3 and E3a (pink highlights), as we might expect, E3 and E3a being the same league in consecutive seasons; but also between I1 and G1 (blue highlights). However, these are mere nuances. The very high correlation throughout the table suggests we could construct a single ‘super model’ of how chances are created that would serve almost as well as any of the 9 league-specific sub-models. A quick and dirty way would be to average, column-wise, the coefficients in the table above. The very great similarity of E3 with E3a coefficients suggests that a one-size-fits-all set of coefficients might also work well, if not in perpetuity, then at least for a few seasons.

This stable and highly portable *model of the many* allows us to clearly identify the *few* who overperform in chance creation. Some, such as Bruno Fernandes, are given every opportunity to express their creativity, and what a happy synergy that can be! Others are not given the opportunities their talents deserve. Yet others, hidden down among the negatives, are players who do not make best use of their opportunities – even though they may create chances, they do not create enough. All are revealed by the model.

John Doyle, Football Analytica