

Development Over Expectation

Who turns recruiting stars into on-field production, and who wastes them

Armaan Sharma

July 2026

Contents

1	Executive summary	2
2	Methodology	4
2.1	Data sources and cohort window	4
2.2	Position scope	4
2.3	The recruit-to-production join	4
2.4	Career-outcome tiers	5
2.5	Transfer attribution	5
2.6	The two-axis design	5
2.7	Expectation baselines	6
2.8	The wepa upgrade and its coverage constraint	6
2.9	The committed_to structural gap	7
3	Findings	8
3.1	Opportunity: who gets more (or fewer) of their signees onto the field	8
3.2	WR production over expectation (WR only, see note)	10
3.3	The two-axis map	12
3.4	Star-tier view: where does production over expectation matter most	15
4	Limitations	17
5	Closing note	19

1 Executive summary

Most programs already know who they signed. Recruiting rankings are public, widely tracked, and treated as a proxy for roster quality. Far fewer programs rigorously measure who they *developed* once those recruits arrived on campus. This report builds an expected-production baseline from recruiting rating and position, then measures which programs beat or trail that baseline, on two axes that are kept deliberately separate: **opportunity** (of the recruits a program signed, what share ever became a meaningful on-field producer at all) and **production over expectation** (among the recruits who did produce, how much they exceeded or trailed the production expected for their rating). That second label is deliberate, not the originally planned “development premium” - as Section 3.2 details, this metric conflates coaching development with offensive scheme, pass volume, target competition, and career length, and a diagnostic check finds those confounds are real enough that the metric should be read as production relative to recruiting pedigree, not as a clean measure of development.

The honest scope, stated up front: at a sample threshold large enough to trust (111 qualifying program-position cells for opportunity, requiring at least 15 recruits per cell), opportunity is measurable across QB, RB, WR, and TE. Production over expectation is a different story: only WR clears the same bar. Nine recruiting classes, a two-season floor to count as a “producer,” and (for QB and RB) the added requirement of clearing the opponent-adjustment metric’s own inclusion threshold combine to leave zero qualifying program-position cells for QB, RB, or TE at $n \geq 15$. That is a finding about data availability, not a flaw in the method, and it is not a threshold we lowered to manufacture more rows. The defensible headline is narrower than the original ambition, but it is real: **WR production over expectation is measurable at the program level, opportunity is measurable more broadly across positions, and the two axes tell meaningfully different stories about the same programs - with the caveat that the production axis is not a clean development signal.**

Three standout findings anchor the two-axis map:

- **Penn State and Texas Tech post WR production well above expectation despite unremarkable opportunity.** Both sit near the middle of the pack on getting signees onto the field at all, but their WRs who do produce substantially outproduce what their recruiting ratings predicted. This is the pattern the project set out to find, but it is not a clean read: Texas Tech in particular ran one of the highest-pass-volume offenses in the qualifying set (see Section 3.2), so scheme is a live alternative explanation alongside coaching development.
- **A cluster of Group of 5 programs (East Carolina, San José State, Texas State) get far more of their signees onto the field as producers than their ratings predict.** This is an opportunity finding, not a production one: it says more about depth chart access than about coaching quality, but it is a real, actionable signal for programs building around lower-rated recruits.
- **Service academies cluster low on opportunity, though not uniformly.** Air Force is below expectation at all four qualifying positions, and Army is below at all four it qualifies for. Navy is below expectation at QB and RB but is actually slightly *above* expectation at WR, so the pattern is real but not monolithic. Where it holds, it should be read in context, not as a criticism: the service academies’ roster model (option offenses, size and commitment requirements, service commitments after graduation) is structurally different from the rest of the sample, and low opportunity rates there likely reflect that structural difference rather than a development failure.

Three structural limitations shape every number in this report and are repeated throughout, not confined to a single section: **program-level opportunity rates are lower bounds**, because roughly half of never-produced recruits have no recorded team commitment in CFBD's database and cannot be attributed to a program; **production-over-expectation figures for QB and RB reflect only players who clear the opponent-adjusted metric's own inclusion bar** (roughly half of producers at those positions), which is a mild survivorship condition inside that axis itself; and **the WR production axis is confounded with offensive scheme, pass volume, and career length**, so it should be read as production relative to recruiting pedigree, not as an isolated measure of coaching development.

2 Methodology

2.1 Data sources and cohort window

Recruiting data (`cfbd_recruiting_player`) and production data (`cfbd_metrics_ppa_players_season`, `cfbd_metrics_wepa_players_passing`, `cfbd_metrics_wepa_players_rushing`, `cfbd_stats_season_player`) come from the CollegeFootballData API via the `cfbfastR` R package. The recruiting cohort covers high school signees from the **2013 to 2021** classes. **Rationale:** production data (PPA) is unreliable before 2013, and 2021 is the latest class whose career fits inside a five-year window ending in 2025, the most recent completed season at the time of this analysis - so every class in the cohort has a genuinely complete career window, not a truncated one. JUCO and Prep School signees were pulled but excluded from the primary cohort. **Rationale:** their prior competitive level is not directly comparable to a four- or five-year high school recruiting rating, so folding them in would mix two different measurement scales.

2.2 Position scope

The primary cohort covers **QB, RB, WR, and TE** - offensive skill positions with individual, publicly available production data. **Rationale:** offensive line has no individual production metric on free data, and defense has no opponent-adjusted value metric at all on this API, only counting stats (see Limitations). Recruiting's own position taxonomy required a mapping step: quarterbacks are labeled "PRO" or "DUAL" (not "QB") and some running backs are labeled "APB." Recruits listed as "ATH" (athlete) were excluded from the primary cohort. **Rationale:** an athlete's eventual position is genuinely undetermined at signing, and assigning one would be a fabricated label, not a recovered one.

2.3 The recruit-to-production join

CFBD's recruiting and production endpoints share an `athlete_id` field, but it is populated for only about half of recruiting records, and even `athlete_id`-to-`athlete_id` matching alone recovers a modest fraction of true matches. A three-tier match was used instead, applied in order:

Table 1: Recruit-to-production join, by match tier

Tier	Rule	Matches
1. Direct <code>athlete_id</code>	Recruiting <code>athlete_id</code> matches a production <code>athlete_id</code> directly	2,918
2. Unique name + position	Exactly one candidate nationally shares the recruit's normalized name and position across all production seasons	2,446
3. Name + position, disambiguated by signing school	Multiple candidates share name and position; accepted only if exactly one candidate's first production season is at the recruit's signing school	26

Rationale for the tiered design: a name-only match risks attributing a different, same-named player's production to the wrong recruit; requiring either a direct id or a school-confirmed identity keeps the false-positive rate low while still recovering the large share of recruits whose recruiting

record lacks an athlete_id. This produced **5,390 matched recruits (50.7%)** out of 10,628 primary-position HS recruits in the cohort window. A further **40** recruits (0.38%) had a genuine name collision with no way to disambiguate and were dropped and counted separately - not silently folded into “never produced.” Recruits with literally no candidate anywhere in the production data (**5,198**, or 49.1% of the cohort) are the “never produced” group and **remain in the cohort** as the backbone of the opportunity axis - dropping them would recreate the exact survivorship bias this project exists to avoid.

2.4 Career-outcome tiers

Each matched recruit is assigned one of three outcome tiers (plus “Never produced” for unmatched recruits), based on a season “qualifying” as a meaningful role - not token action. PPA’s own play-count field was unpopulated in this data version, so a usage-volume proxy was built from counting stats instead: QB needs 100+ combined pass/rush attempts in a season, RB needs 40+ touches, WR needs 15+ catches, TE needs 10+ catches (TE’s lower bar reflects the position’s naturally smaller target share, not a looser standard). **Rationale:** these thresholds were set from the empirical usage distribution of matched players, not chosen to hit a target tier size.

Table 2: Career-outcome tier distribution, full cohort

Career-outcome tier	N	% of cohort
Never produced	5198	49.1
Rotational contributor	2958	27.9
Multi-year producer	1945	18.4
Star producer	487	4.6

Multi-year producer requires 2+ qualifying seasons; Star producer requires 2+ qualifying seasons **and** career production in the top 20% of multi-year producers at that position. As a sanity check, the ever-produced rate rises monotonically with recruiting stars (2-star: 32.9%, 3-star: 56.9%, 4-star: 86.6%, 5-star: 94.5%) - confirming the join is capturing real signal, not noise.

2.5 Transfer attribution

Each producing season is attributed to the team the player actually played for that season, not the team he signed with. A player counts as “non-transfer” if every production season is at his signing school, and “transfer” if any season is elsewhere. **45.6%** of producers are non-transfers, **31.4%** transferred at some point, and the remainder have an unknown signing school (see below). **Rationale:** attributing a transfer’s production to his signing school would overstate that program’s development, and attributing it to no one would understate the transfer destination’s real production - crediting the actual team played for is the only defensible rule.

2.6 The two-axis design

Opportunity and development are computed and reported completely separately, and never collapsed into one score. **Rationale:** a program that stockpiles blue-chip recruits and buries most of

them looks very different from a program that turns unheralded recruits into producers, but a single blended “development score” would erase that difference. Measuring development only among players who saw the field is survivorship-biased on its own - deep blue-blood rosters bury more talent, so their producers are a more selected group - which is exactly why opportunity (who plays at all) has to be its own axis, with the full recruit population, including washouts, in the denominator.

2.7 Expectation baselines

Two baselines are fit as smooth (GAM) functions of continuous recruiting rating within position group, controlling for signing class year:

- **Opportunity baseline:** a logistic GAM predicting P(ever produced | rating, position, year), fit on the full cohort. **Rationale:** a flexible smooth term captures a non-linear rating-to-opportunity relationship better than treating stars as a simple linear predictor.
- **Development baseline:** a Gaussian GAM predicting expected production level given rating, position, and year, fit only on players who produced (the resulting premium is explicitly conditional on producing). This is the statistical name for the baseline; as Section 3.2 details, the WR premium it produces should be read as production over recruiting pedigree, not as an isolated measure of coaching development. **Rationale for the Gaussian family:** an earlier version used a robust/heavy-tailed family (`scat()`) to handle the long right tail of production totals, but a calibration check (mean of actual vs. mean of fitted, by position) showed it systematically underfit the true mean for WR (fitted 35.0 vs. actual 56.8) - a robust estimator targets a different central tendency than the mean, and “premium” is only meaningful relative to the honest average. Gaussian regression targets the true conditional mean by construction; the same calibration check now agrees with the actual mean to floating-point precision for all four positions.

The development baseline is fit twice: with and without a strength-of-schedule control (a team’s SP+ overall rating, averaged across a player’s production seasons). The with-SOS version is the primary reported premium; the without-SOS version is shown alongside on every leaderboard as a sensitivity check, not hidden in an appendix. The two versions correlate at 0.986 - the SOS control adds real explanatory power (adjusted R-squared rises from 0.277 to 0.325) without drastically reshuffling which programs look over- or under-performing.

2.8 The wepa upgrade and its coverage constraint

Opponent-adjusted production (`wepa`) became available partway through this project after a CFBD API tier upgrade. It is now the **primary** production metric for QB (passing + rushing `wepa` summed) and RB (rushing `wepa` only) in Phase 4 and 5. WR and TE remain on raw `total_PPA_all`. **Rationale:** `wepa` has no receiving-play adjustment on this API, so it cannot cover WR/TE’s overwhelmingly receiving-based value, and RB’s own receiving contribution is similarly not captured - RB production-over-expectation figures are a rushing-value premium, not a full-value one. `wepa` also carries its own (stricter, undocumented) minimum-plays inclusion bar: only **52.4%** of QB producers and **52.8%** of RB producers clear it, versus 100% of WR/TE (raw PPA has no such filter). This is a mild survivorship condition inside the production axis itself: QB and RB production-over-expectation figures are conditional on both producing and clearing `wepa`’s bar, and are not presented as equivalent in coverage to WR/TE anywhere in this report.

2.9 The committed_to structural gap

Roughly 36% of the cohort has no recorded signing school (`committed_to`) in CFBD's recruiting data, rising to about 49% among never-produced recruits specifically. A recovery attempt re-pulled recruiting data team by team (`cfbd_recruiting_player(year, team = X)`, which returns only recruits confirmed committed to team X) across all FBS programs and years, on the theory that the aggregate endpoint might be missing associations the team-scoped endpoint could recover. It recovered only 5 additional records. **Why:** a direct test showed that recruits who *do* have a known signing school appear in the team-scoped data 80.9% of the time (confirming the team-scoped endpoint is a faithful subset of the same data), while recruits *missing* a signing school appear there only 0.1% of the time. CFBD's database itself never recorded a team association for these recruits - there is no hidden data a team-by-team query can surface. This is a structural completeness limit in the source data, not a bug in extraction, and is treated as permanent: these recruits are excluded from program-level cells but remain in every overall/aggregate rate.

3 Findings

3.1 Opportunity: who gets more (or fewer) of their signees onto the field

Program-level opportunity rates are lower bounds. Roughly 50% of never-produced recruits have no recorded team commitment in CFBD's database and are excluded from program-level denominators.

Table 3: Opportunity leaderboard, top 10 qualifying cells, ALL positions pooled (n >= 15)

Program	Position	N	Actual rate	Expected rate	Premium
East Carolina	WR	16	81%	42%	+39pp
San José State	WR	22	82%	43%	+39pp
Texas State	WR	15	80%	43%	+37pp
Kansas	WR	19	79%	48%	+31pp
Charlotte	WR	19	68%	40%	+29pp
Bowling Green	RB	15	80%	51%	+29pp
Kent State	WR	16	69%	40%	+29pp
Houston	WR	17	82%	54%	+28pp
Temple	WR	24	71%	44%	+27pp
Massachusetts	WR	16	69%	43%	+25pp

Table 4: Opportunity leaderboard, bottom 10 qualifying cells, ALL positions pooled (n >= 15)

Program	Position	N	Actual rate	Expected rate	Premium
Air Force	RB	24	17%	40%	-24pp
Arizona	WR	17	35%	55%	-20pp
Air Force	TE	26	15%	34%	-19pp
Minnesota	WR	15	47%	59%	-13pp
Army	TE	15	20%	32%	-12pp
Notre Dame	WR	18	72%	82%	-10pp
Northern Illinois	RB	16	38%	47%	-10pp
Air Force	WR	29	28%	37%	-9pp
Army	QB	20	25%	33%	-8pp
Air Force	QB	19	26%	34%	-7pp

Tables 3 and 4 pool all four positions together - a program can appear more than once if it has qualifying cells at multiple positions (Air Force does, four times). Figure 1's opportunity axis later in this report also blends positions, but its production axis is WR only; Tables 5 and 6 and Figure 3 below are WR-specific. Position composition differs across these views, and each caption states which positions it covers.

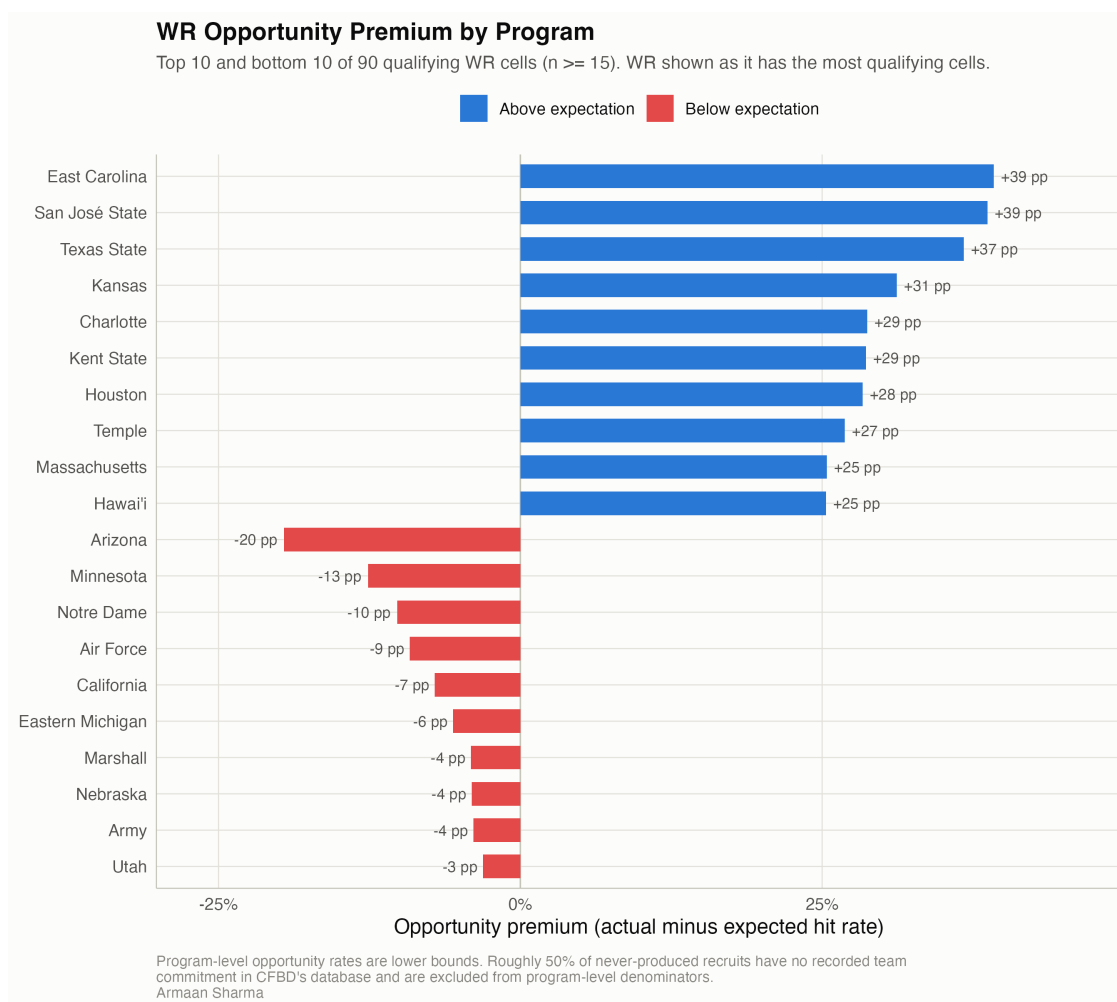


Figure 1: WR opportunity premium, top and bottom 10 of 90 qualifying cells.

Of 889 program-position cells, 111 clear the $n \geq 15$ bar (QB: 3, RB: 16, WR: 90, TE: 2) - WR and RB have enough qualifying volume for a broad opportunity read; QB and TE do not, and their leaderboard entries should be treated as suggestive at best.

A contextual note on the service academies. The pattern is real but not uniform. Air Force is below expectation at all four positions it qualifies for (QB -7pp, RB -24pp, WR -9pp, TE -19pp), and Army is below at all four it qualifies for. **Navy's pattern is mixed:** below expectation at QB (-7pp) and RB (-5pp), but actually *above* expectation at WR (+6pp). This is flagged here as context, not criticism: these programs run option-heavy offenses with different positional usage patterns, recruit under size and commitment constraints the rest of the sample does not face, and commit players to a post-graduation service obligation that shapes roster construction in ways unrelated to coaching or development quality. Where the pattern holds, it most likely reflects this structural difference rather than a development failure.

3.2 WR production over expectation (WR only, see note)

QB/RB development n reflects only players clearing wepa's inclusion threshold (~52-53% of producers) - a mild survivorship element within the development axis itself. Not equivalent in coverage to WR/TE (100% raw-PPA coverage).

This metric is not a clean measure of coaching development, and that is worth stating here, in the findings, not only in the limitations section. WR production over expectation is raw production relative to a rating-and-schedule-based baseline. It does not, and cannot, isolate coaching development from three other real drivers: offensive scheme and pass volume (a pass-heavy program mechanically produces more raw receiving value per WR, independent of coaching quality), competition for targets within a loaded receiving room, and career length (a producer who leaves early for the NFL has fewer seasons counted toward his total). The strength-of-schedule control used elsewhere in this report addresses none of these three. To find out how much they matter in practice, the with-SOS premium for each of the 25 qualifying programs was correlated against three program-level proxies:

Table 5: WR production over expectation vs. three confound proxies, all 25 qualifying programs

Program	N	Premium	Pass att/szn	Avg seasons	Avg rating
Penn State	18	26.7	375	3.78	0.902
Texas Tech	15	25.9	512	3.73	0.857
Oklahoma	17	15.8	366	3.35	0.922
Washington	15	12.7	367	3.33	0.904
Ole Miss	18	11.2	398	3.44	0.905
LSU	25	5.8	365	3.04	0.925
Baylor	16	3.3	411	3.62	0.889
Ohio State	23	2.9	365	3.35	0.944
San José State	18	1.2	415	3.22	0.799
Alabama	20	-0.3	375	3.55	0.944
USC	16	-1.6	420	3.44	0.931
Auburn	17	-1.6	325	3.76	0.911
Temple	17	-3.3	398	3.35	0.804
UCLA	15	-3.5	391	3.60	0.888

Miami	17	-3.6	393	3.59	0.909
Clemson	18	-5.2	441	3.50	0.937
Kansas	15	-9.9	362	3.27	0.815
Michigan	17	-11.3	322	3.65	0.890
Florida State	20	-16.3	376	3.75	0.919
Georgia	20	-17.9	359	3.40	0.909
Florida	16	-20.9	358	3.81	0.926
South Carolina	15	-21.1	369	3.00	0.876
Tennessee	16	-23.7	350	3.19	0.896
Texas	22	-26.0	392	3.09	0.913
Texas A&M	19	-28.9	394	3.26	0.925

Pass att/szn = avg. team pass attempts per season. Avg seasons = avg. seasons played by that program's WR producers. Avg rating = avg. WR recruiting rating signed.

Table 6: Confound diagnostic: correlation of WR premium with three proxies (n = 25 programs)

Confound proxy	Correlation with premium
Team pass volume (avg. season pass attempts)	0.358
Career length (avg. seasons played by WR producers)	0.312
Avg. WR recruiting rating signed	-0.079

The pass-volume correlation ($r = 0.358$) is the largest of the three and is not trivial: **Texas Tech**, the second-highest premium in the qualifying set, also ran the highest pass volume of any qualifying program in this window (512 attempts per season, versus a qualifying-set average in the mid-to-high 300s). That is exactly the pattern the diagnostic was built to catch, and it means Texas Tech's premium is at least partly a scheme effect, not purely a coaching one. The correlation is moderate and only marginally significant at this sample size ($n = 25$; a two-sided test gives p around 0.08), so it should be read as suggestive evidence of a real confound, not proof that any single program's premium is entirely scheme-driven. The career-length correlation ($r = 0.312$) runs in the opposite direction from the "early NFL departure inflates value" concern: programs whose WR producers stick around *longer* show somewhat *higher* premiums, not lower, which is more consistent with sustained production accumulating across a longer career than with a few short, explosive careers driving the number up. The recruiting-rating correlation ($r = -0.079$) is close to zero, which is expected since the baseline already conditions on rating.

The bottom line: this metric should be read as production over recruiting pedigree, not as a clean measure of player development. The programs at the top of the leaderboard are producing WR value beyond what their recruits' ratings alone predict; whether that reflects coaching, scheme, or some mix of both is not something this metric can separate on its own.

Table 7: WR Production Over Expectation by Program, top 10 of 25 qualifying programs ($n \geq 15$)

Program	N	With SOS (primary)	Without SOS (sensitivity)	Diverges
Penn State	18	26.7	29.8	
Texas Tech	15	25.9	24.4	

Oklahoma	17	15.8	18.4	
Washington	15	12.7	15.7	
Ole Miss	18	11.2	10.8	
LSU	25	5.8	5.3	
Baylor	16	3.3	6.8	
Ohio State	23	2.9	5.9	
San José State	18	1.2	7.7	Yes
Alabama	20	-0.3	1.9	

Table 8: WR Production Over Expectation by Program, bottom 10 of 25 qualifying programs (n >= 15)

Program	N	With SOS (primary)	Without SOS (sensitivity)	Diverges
Texas A&M	19	-28.9	-27.7	
Texas	22	-26.0	-24.7	
Tennessee	16	-23.7	-21.8	
South Carolina	15	-21.1	-23.7	
Florida	16	-20.9	-19.2	
Georgia	20	-17.9	-16.7	
Florida State	20	-16.3	-14.4	
Michigan	17	-11.3	-7.2	
Kansas	15	-9.9	-8.4	
Clemson	18	-5.2	-2.7	

Divergence cases warranting extra caution. Two programs in the qualifying set show a meaningfully different read depending on whether the strength-of-schedule control is applied. **Temple** shows a without-SOS premium of +6.9 but a with-SOS premium of -3.3 - once its schedule strength is accounted for, a seemingly average-to-good WR production record looks below expectation instead. **San José State** shows the reverse pattern less dramatically: +7.7 without SOS versus +1.2 with SOS - still positive, but meaningfully smaller once schedule is controlled for. Neither program appears in Figure 2's top/bottom 10 by the primary with-SOS metric (Temple's -3.3 and San José State's +1.2 both sit inside the middle of the 25-program distribution), but both are worth flagging in prose precisely because the choice of metric changes the story.

3.3 The two-axis map

Only the 25 programs qualifying on **both** axes (n >= 15 for opportunity and for WR production) are plotted; a program with a qualifying score on only one axis is not given a half-valid coordinate. The x-axis blends whichever positions a program qualifies on for opportunity; the y-axis is WR only. 68 programs qualify on opportunity only and are listed in Table 9 instead of plotted.

The map's quadrant labels use "production," not "development," for the same reason the rest of this section does: the y-axis cannot separate coaching development from scheme, target competition, and career length. Five programs anchor the map's interesting quadrants:

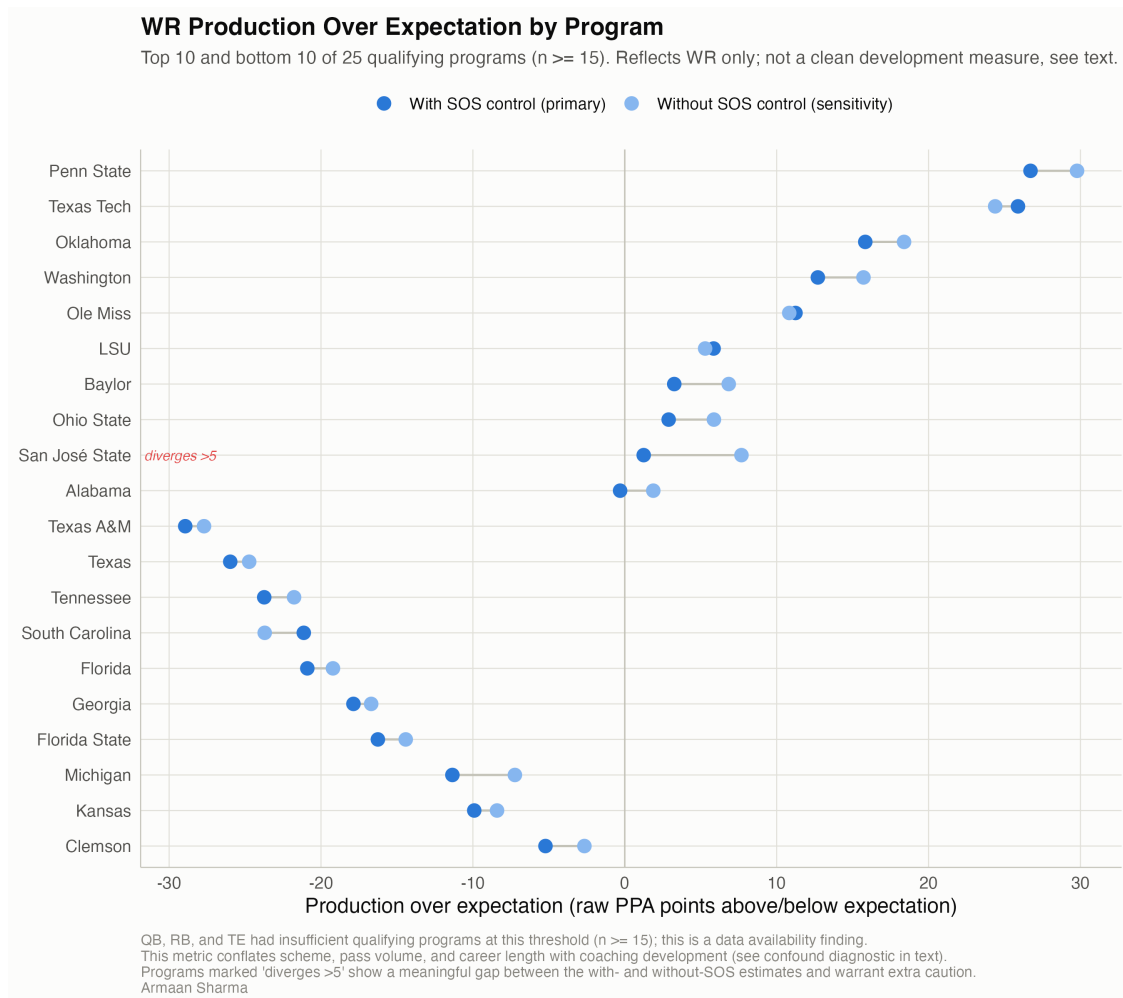


Figure 2: WR production over expectation, with and without the SOS control, top and bottom 10 of 25 qualifying programs.

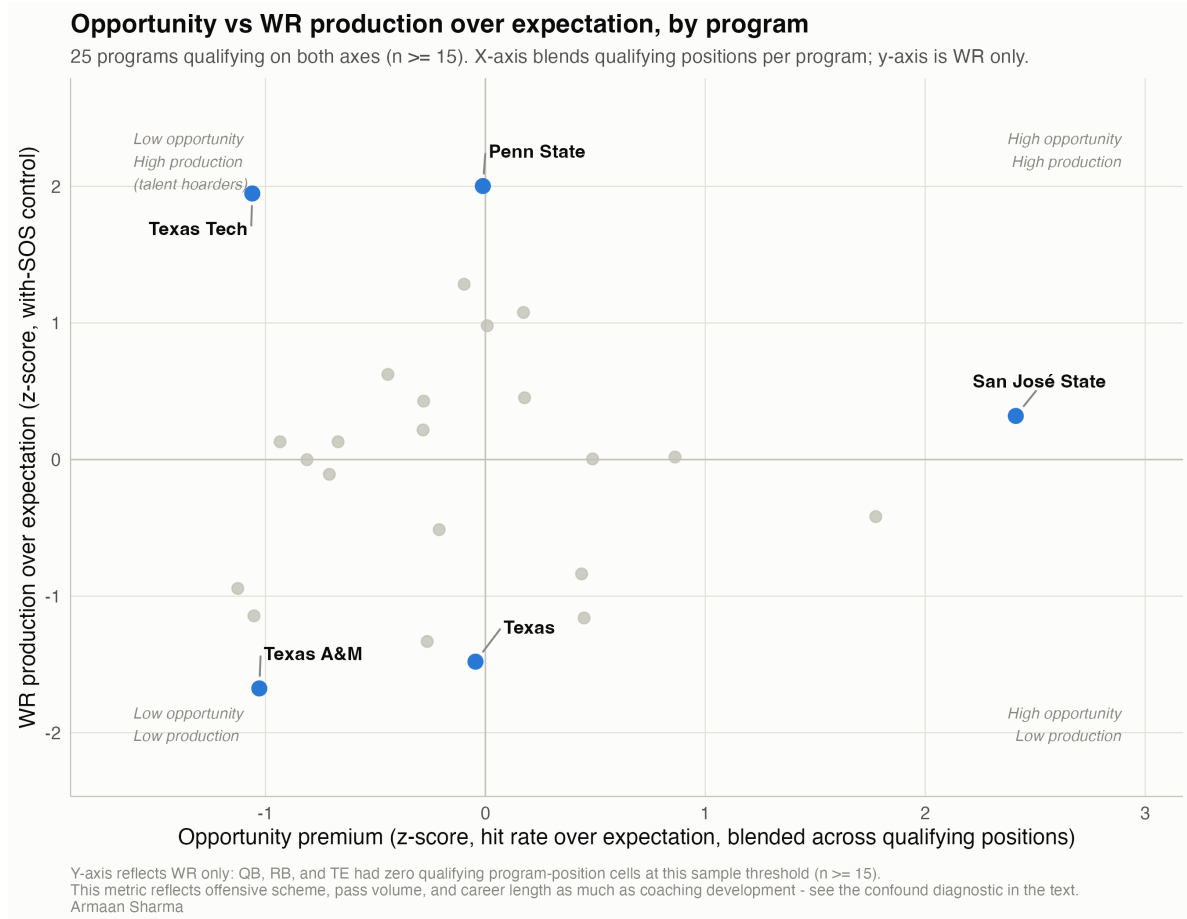


Figure 3: Opportunity vs WR production over expectation, 25 programs qualifying on both axes.

- **Penn State** and **Texas Tech** sit in the low-to-neutral opportunity, high production quadrant - the map's "talent hoarders" label, describing a roster-construction pattern (few signees reach the field, but those who do substantially outproduce expectation) rather than a coaching claim. That distinction matters here specifically: the confound diagnostic above found Texas Tech ran the highest pass volume in the qualifying set, so scheme is a live alternative explanation for at least part of its premium, alongside whatever coaching effect exists. Penn State's pass volume is close to the qualifying-set average, so scheme is a weaker alternative explanation for its premium specifically.
- **San José State** sits in the high opportunity, modest production quadrant - it gets an unusually large share of its WR signees onto the field as producers, but those producers perform closer to (rather than above) expectation once they arrive.
- **Texas A&M** and **Texas** anchor the bottom of the production axis with below-average opportunity as well - by these two metrics, both the hit rate and the premium on WRs who do hit lag their recruiting investment.

Table 9: Programs qualifying on one axis only (15 of 68 shown; not plotted on the map)

Program	Qualifies on	Recruits (opportunity denom.)	Producers (development denom.)
Air Force	opportunity only	98	17
Army	opportunity only	82	11
Navy	opportunity only	74	16
Miami (OH)	opportunity only	67	27
Northern Illinois	opportunity only	66	25
Old Dominion	opportunity only	59	21
Marshall	opportunity only	58	22
Wyoming	opportunity only	58	27
Bowling Green	opportunity only	57	27
Western Michigan	opportunity only	56	24
San Diego State	opportunity only	55	27
Central Michigan	opportunity only	54	21
Coastal Carolina	opportunity only	53	23
Eastern Michigan	opportunity only	52	17
Florida International	opportunity only	52	19

3.4 Star-tier view: where does production over expectation matter most

A lower, explicitly disclosed sample bar applies only to this view. Every other leaderboard and the map in this report uses $n \geq 15$. Splitting the WR qualifying set further by star tier leaves too few programs per cell at that bar to compute a meaningful spread (raising it to $n \geq 15$ here would leave a single tier - 4-star - with five programs and zero programs everywhere else, which is not a usable view). This section uses **$n \geq 5$ producers per program-star cell** instead, which is noisier and should be read with more caution than the headline leaderboards above.

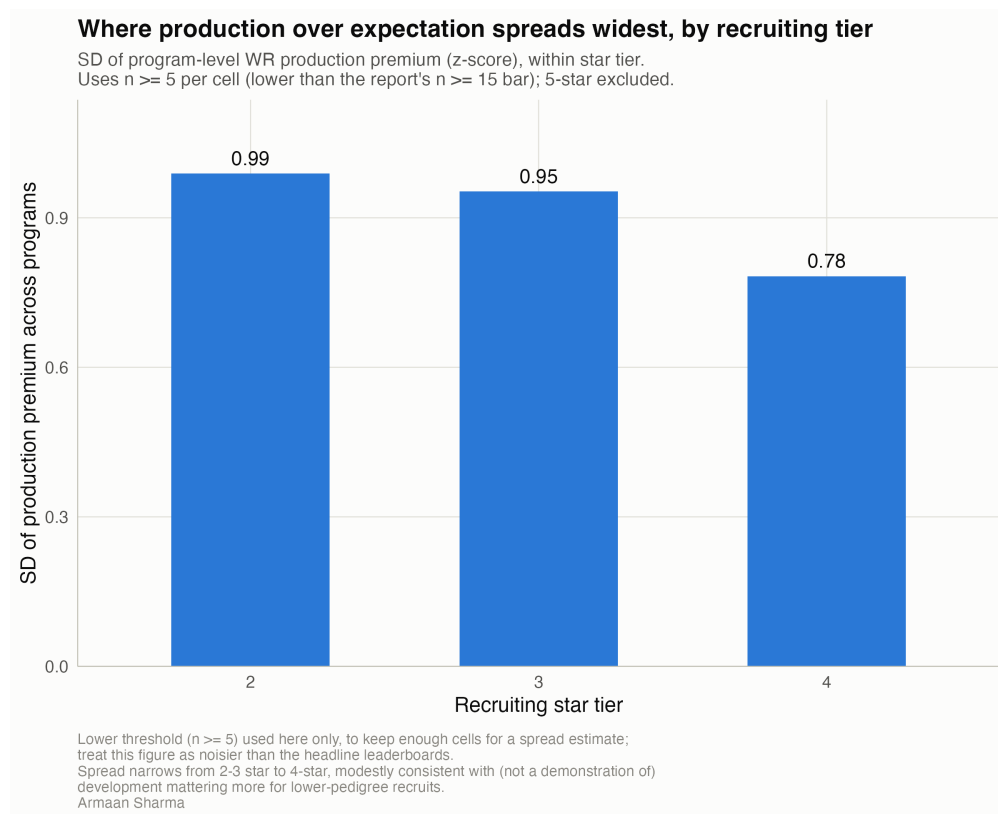


Figure 4: Standard deviation of program-level WR production-over-expectation premium, by recruiting star tier ($n \geq 5$, a lower bar than the rest of the report).

Table 10: Program-level WR production-over-expectation spread, by star tier (n >= 5 per cell, 5-star excluded, n too small)

Stars	Programs	SD across programs	Min	Max
2	35	0.99	-2.27	3.05
3	86	0.95	-1.98	3.59
4	29	0.78	-1.68	1.91

The spread across programs narrows from 2-star and 3-star recruits down to 4-star recruits (0.99 and 0.95 versus 0.78). This is **modestly consistent with** development mattering more for lower-pedigree recruits - it does **not demonstrate** it, both because of the confounds discussed above and because this specific view runs on a noisier n >= 5 bar. The 2-star versus 3-star difference (0.99 vs. 0.95) is small enough that it should not be read as a meaningful effect on its own; the clearer signal, if there is one, is the narrower spread at 4-star. Five-star recruits are excluded entirely because too few programs have enough five-star WR producers to compute a meaningful program-level spread even at the lower n >= 5 bar.

4 Limitations

- **Program-level opportunity rates are lower bounds.** Roughly 36% of the cohort overall, and about 50% of never-produced recruits specifically, have no recorded team commitment in CFBD's database and are excluded from program-level denominators (they remain in every overall/aggregate rate). This is a structural gap in the source data, confirmed by a team-by-team recovery attempt that found the missing association simply does not exist in CFBD's system - not a bug in this project's extraction. Programs with disproportionately many unattributable washouts will look better on opportunity than they truly are.
- **The WR production-over-expectation metric conflates coaching development with scheme, target competition, and career length, and the strength-of-schedule control addresses none of these.** A diagnostic check (Section 3.2) found a moderate, marginally significant correlation (r = 0.36, p around 0.08) between a program's premium and its team pass volume - concretely, Texas Tech's premium coincides with the highest pass volume in the qualifying set. This metric should be read as production relative to recruiting pedigree, not as an isolated measure of player development.
- **Production-over-expectation rankings reflect WR only at the current sample threshold (n = 15).** QB, RB, and TE had zero qualifying program-position cells at this threshold, driven by a combination of the 9-class cohort window, the 2-season "producer" floor, and (for QB/RB) wepa's own coverage cut. The threshold was not lowered to manufacture more cells: doing so would introduce noise that undermines the reliability of the estimates being reported. The star-tier view (Section 3.4) uses a lower n >= 5 bar, disclosed explicitly there, to keep enough cells for a spread estimate.
- **QB and RB production-over-expectation figures carry a mild survivorship condition of their own.** Only about 52-53% of QB and RB producers clear wepa's inclusion bar; their figures are conditional on both producing and clearing that bar, and are not equivalent in coverage to WR/TE (100% raw-PPA coverage, no additional filter).

- **Offensive line is excluded entirely** - no individual production metric exists for OL on free data.
- **Defense is out of scope for this report.** It has no opponent-adjusted value metric on this API, only counting stats (tackles, TFL, sacks, interceptions, pass breakups), which are not a value or efficiency measure and would not be comparable to the offensive skill-position analysis here.
- **WR and TE production is raw PPA, not opponent-adjusted.** wepa has no receiving-play adjustment on this API. WR/TE figures reflect raw value against the field played, not schedule-adjusted value - this is part of why the metric is reported as production over expectation rather than a clean development premium.
- **The cohort is 2013-2021 high school recruits only.** JUCO and Prep School signees were pulled but excluded from the primary analysis; their prior competitive level is not on the same measurement scale as a multi-year high school recruiting rating.
- **All findings in this report are associational, not causal.** A program's opportunity or production-over-expectation premium reflects the observed relationship between recruiting rating and outcome for its signees; it does not isolate coaching effect from scheme fit, strength staff, playing-time politics, roster context, or any other mechanism, and should not be read as proof that a specific intervention caused a specific outcome.

5 Closing note

This report is a quantitative supplement to scouting, not a replacement for it. It quantifies base rates and premiums across a large sample of recruits and programs; it does not evaluate any individual recruit's specific development trajectory, and it should not be used as the sole basis for a roster-construction decision about any one player or coaching staff. The value of this kind of analysis is in surfacing patterns - a program producing more than its recruiting pedigree predicts, a program that outpunches its recruiting weight on opportunity, an opportunity pattern that deserves a second look - that a staff can then investigate with the scouting, film, and internal context this report does not have access to.