# Mid-Major to High-Major Transfer Translation Rates

A quantitative supplement to scouting, not a replacement

Armaan Sharma

July 03, 2026

## Contents



## 1 Executive Summary

Moving from a mid-major program to a high-major program is associated with a cost to both **role** and **efficiency**, and the version of this question almost anyone would answer from raw box scores misses the second half of that cost entirely. Restricted to players who held a genuine rotation role in

their first high-major season (the "strict" sample, n = 248), adjusted offensive rating rises slightly after an upward transfer (+2.6) and the relationship between a bigger level jump and post-transfer efficiency is statistically flat. But 48.8% of upward transfers in the full matched cohort (n = 549) never reach rotation minutes at the new level at all, a washout rate that ranges from roughly 32% for secondary creators and combo guards (who clear rotation about two-thirds of the time) to roughly 72% for back-to-the-basket, rim-running bigs. Once the partial-washout players who cleared some bar short of a full role are restored to the sample ("broad", n = 382), the picture reverses: efficiency falls (-1.5 adjusted ORtg on average) and a bigger level jump becomes a real, statistically significant predictor of a larger efficiency decline. The strict-sample numbers are not wrong, but they describe only the players who survived the jump, and survivorship is exactly the wrong lens for projecting a specific transfer target who has not yet proven he can hold a role. Ball-handling and shot-creation skill is the most reliable thing that translates upward; a traditional low-post role is the least reliable.

# 2 Methodology

## 2.1 Data sources

Player-season stats and team-season ratings come from `cbbdata` (Barttorvik data, 2018-2025 in this run). No working transfer-portal data source exists in `cbbdata` or in `toRvik` (the package specified in the original brief) as of this analysis: `toRvik` was archived by its maintainer in September 2024, and the backend API its `transfer_portal()` function depends on is no longer live. Barttorvik's own transfers page is also not scrapable behind a bot-verification wall.

Transfers were instead **derived** from `hoopR` (ESPN) player box-score rosters: if a player's primary team changes between two seasons, that is treated as a transfer. Team identifiers are mapped between ESPN and Barttorvik team names using `cbbdata::cbd_teams()`, which provides an exact numeric join (`espn_id`), not a fuzzy one. Player identity across the pre- and post-transfer season is joined by normalized name matching within the correct team-season roster (a small, known set of names, not a national search), after stripping generational suffixes and punctuation. This method has one structural blind spot: a player who transfers and never appears in a single game box score at the new school (no game dressed at all) is invisible to it entirely. This is addressed directly in Limitations.

## 2.2 The detection funnel

Every filtering step from raw box-score rows to the final analysis sample is tracked explicitly so the sample size is auditable rather than a single before/after number.

Table 1: Transfer detection and cohort-assembly funnel

| Stage | n |
|---|---|
| hoopR athlete-season-team rows (incl. dual-team seasons) | 109946 |
| athlete-seasons after primary-team resolution | 109536 |
| consecutive-season (t -> t+1) team changes detected | 8208 |
| both teams mapped to a Barttorvik team name | 6942 |
| both team-seasons have a Barttorvik rating (D1 both sides) | 4023 |

| | |
|---|---|
| classified mid/high (both seasons) | 4023 |
| upward (mid -> high) moves only | 586 |
| deduped on (athlete_id, pre_season, post_season) | 586 |
| matched to pre-transfer (mid-major) season stats | 549 |
| final upward cohort (pre matched; post matched or labeled washout) | 549 |
| analysis-eligible: rotation minutes BOTH seasons (STRICT) | 248 |
| analysis-eligible: BROAD (rotation + limited_minutes post, rotation pre) | 382 |

### 2.3 Level classification

Team-seasons are classified mid-major versus high-major using Barttorvik's `barthag` rank, with high-major defined as the top 75 teams by `barthag` in that season (a threshold intended for sensitivity testing against 60 and 90 in future work; this run uses 75 throughout). Level is also carried as a **continuous** measure: jump size is defined as the destination team's `barthag` rating minus the origin team's `barthag` rating at the time of the move, not a rank difference. This continuous measure is what feeds the regression models below.

### 2.4 Archetype clustering

Players are clustered into six role archetypes using k-means (k chosen by silhouette score across k = 4 to 6; k = 6 had the best average silhouette, 0.369) on **pre-transfer, mid-major** profiles only: usage rate, assist rate, offensive and defensive rebound rate, shot-location share (rim/mid-range/three-point), and Barttorvik's own estimated position, one-hot encoded. Clustering is fit on the pre-transfer profile only because archetype is meant to describe a player's role identity coming in, not an outcome of the move itself. Because the clustering features are entirely offensive (no steal, block, or defensive rating inputs), archetype labels describe offensive role only; a label like "3-and-D wing" would overclaim a defensive judgment the data does not support, so labels are offense-only throughout (e.g. "low-usage spot-up wing").

Table 2: Archetypes, sample sizes, and washout rate

| Archetype | n (full cohort) | Washout rate |
|---|---|---|
| Rim-running big | 72 | 72.2% |
| Stretch big | 43 | 62.8% |
| Low-usage spot-up wing | 128 | 56.2% |
| Versatile forward (inside-out) | 109 | 45.0% |
| High-usage lead guard / primary creator | 82 | 37.8% |
| Secondary creator / combo guard | 115 | 32.2% |

### 2.5 The usage confound, and why it is treated in two separate places

A player's role mechanically shrinks when he joins a more talented team, and that role compression is entangled with any honest measure of translation. This project separates the confound in two distinct ways rather than one:

**Opponent adjustment (cross-level comparisons).** Barttorvik's `adj_oe` (adjusted offensive rating) is opponent- and schedule-adjusted: a player's raw individual offensive rating is scaled by the ratio of Division I average defensive efficiency to the average defensive efficiency of the opponents he actually faced. This is confirmed directly from Barttorvik's own published methodology. `adj_oe` is therefore the primary efficiency metric used whenever a mid-major season is compared to a high-major season (Section 3.2 and the regression models), since it is the metric that nets out the rise in competition level. TS% and eFG% are reported alongside as a cross-check but are **not** opponent-adjusted, and are treated as descriptive supporting evidence only, never as the primary cross-level comparison.

**Usage adjustment (Section 3.3, the usage-efficiency curve).** Barttorvik's `adj_oe` already has an internal usage-normalization step baked in (applied before the opponent adjustment), which flattens the very usage-efficiency tradeoff that Section 3.3 is designed to model. Using `adj_oe` to build that curve would partially double-adjust for usage. Section 3.3 therefore uses **raw, unadjusted individual ORtg** to fit an expected-efficiency-given-usage curve on a non-transfer high-major population, and compares each transfer's actual post-transfer raw ORtg against that curve. Because the comparison happens entirely within one level (high-major, post-transfer season), no cross-level opponent adjustment is needed at that step; that concern is specific to comparisons across levels, which the `adj_oe`-based sections already handle. The baseline population explicitly excludes the transfer cohort's own post-transfer rows, so transfers are never partly measured against themselves, and baseline residuals were checked for a trend across the usage range where transfers land (18-22%); the observed drift was -0.05 adjusted ORtg points across that band, well inside a 1-point tolerance, so the curve was used as fit.

### 2.6 Total versus direct effect in the regression

The regression models (Section 3.4) also separate two different questions that are easy to conflate. Jump size (a bigger level jump) mechanically causes usage compression, so usage change is a **mediator** of jump size's effect on post-transfer efficiency, not an independent confounder. Controlling for it answers a different question than leaving it out. Two versions of each model are reported: a **total effect** model (jump size only, no usage-change control), which is the headline "what does jumping up cost overall" number, and a **direct effect** model (jump size with usage change held fixed), which isolates the part of that cost that does not run through losing usage. Points per 40 minutes is not regressed at all: with usage change controlled it becomes close to mechanical (points per 40 is largely a function of usage), so it is reported only as a descriptive rate in Section 3.2, not as a regression outcome.

All regression and correlational language in this report is **associational, not causal**. This is an observational cohort of players who selected into transferring, not an experiment; "moving up is associated with X," never "moving up causes X."

# 3 Findings

## 3.1 Washout rate: the finding the strict sample cannot show

Of the 549 upward transfers with a confirmed mid-major baseline, only 51.2% reach rotation minutes (15+ games, 40%+ minutes rate) at the new school in their first season. This is a **lower bound**: it counts only players detected via a hoopR box-score appearance (including a zero-minute, dressed-but-did-not-play row); a player who transfers and never appears in a single box score at the new school is invisible to the pipeline entirely and is not counted here at all.

Table 3: Post-transfer outcome, full matched cohort (n = 549)

| Post-transfer status | n | % of cohort |
|---|---|---|
| Reached rotation | 281 | 51.2% |
| Limited minutes | 189 | 34.4% |
| No production | 79 | 14.4% |

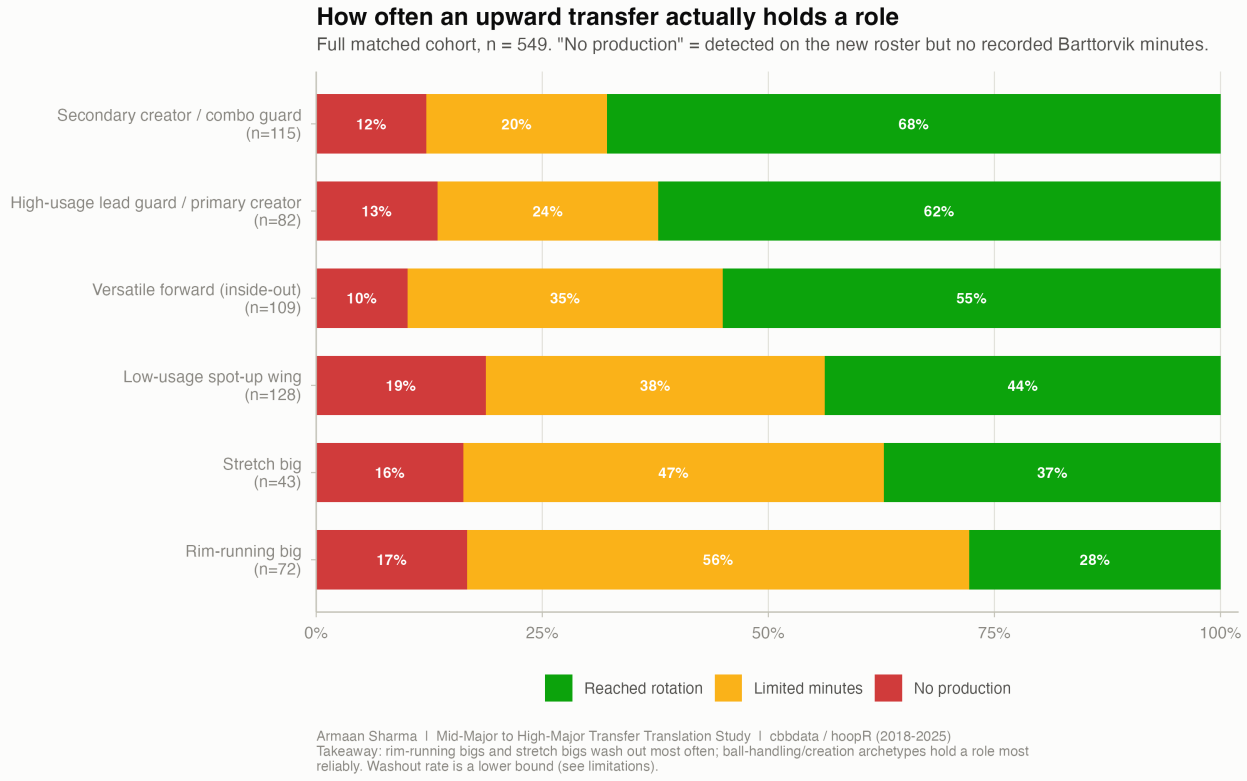


Figure 1: Washout tier by archetype, full matched cohort.

The washout gradient by archetype is wide and directionally intuitive: ball-handling and shot-creation skill (secondary creators, lead guards) survives the jump most reliably, while a traditional low-post role washes out most often, plausibly because those roles depend most on a specific physical/schematic fit that does not transfer as cleanly as perimeter shot-creation.

### 3.2 Strict versus broad: the translation-rate reversal

**Every rate below is labeled strict (rotation-eligible only, n = 248) or broad (rotation plus limited-minutes, n = 382) throughout this report, and should never be read without that label.** The strict sample answers "how did the players who held a role perform." The broad sample restores the partial-washout players a strict minutes filter is specifically designed to hide, and is closer to the honest answer to "how does an upward transfer perform."

Table 4: Raw before/after, all archetypes pooled, strict vs broad (n = 248 strict, 382 broad)

| Metric | Strict Pre | Strict Post | Strict Chg | Broad Pre | Broad Post | Broad Chg |
|---|---|---|---|---|---|---|
| Usage (%) | 24.31 | 19.78 | -4.53 | 23.47 | 18.86 | -4.61 |
| Adj. ORtg | 110.46 | 113.03 | +2.58 | 108.85 | 107.39 | -1.46 |
| TS% | 54.70 | 55.06 | +0.37 | 54.81 | 53.45 | -1.36 |
| Pts/40 | 18.15 | 15.04 | -3.11 | 17.55 | 13.77 | -3.77 |

Usage compression is nearly identical in both samples (roughly -4.5 to -4.6 points): the two samples disagree about efficiency, not about role loss. In the strict sample adjusted ORtg rises 2.58 points (110.46 to 113.03); in the broad sample it falls 1.46 points (108.85 to 107.39). TS% follows the same pattern (+0.37 strict, -1.36 broad) but is not opponent-adjusted, so it is reported as a cross-check, not as the primary read.

### 3.3 Usage-adjusted translation: do transfers beat the efficiency expected at their new usage?

Built on raw ORtg against a non-transfer high-major baseline curve (Methodology, above), not on the opponent-and-usage-adjusted `adj_oe`.

Table 5: Usage-adjusted over/under-performance vs. high-major expectation

| Sample | n | Mean residual | % beat | % meet | % fall short |
|---|---|---|---|---|---|
| Strict | 248 | -0.01 | 44% | 13% | 43% |
| Broad | 382 | -4.36 | 35% | 12% | 53% |

Even in the strict sample, the average transfer is at best roughly neutral against the efficiency a non-transfer high-major player would be expected to produce at the same usage (mean residual -0.01). In the broad sample the average transfer clearly falls short of that baseline (mean residual -4.36). Conditional on holding a role, upward transfers are not outperforming their new level; they are approximately meeting it at best, and the fuller population falls short of it.

### 3.4 Regression: the translation discount, in interpretable units

Jump size is barthag rating delta, which ranges roughly 0.06 to 0.77 in this sample (interquartile range 0.2426). A raw coefficient in "adjusted ORtg points per 1.0 barthag point" is not a jump any

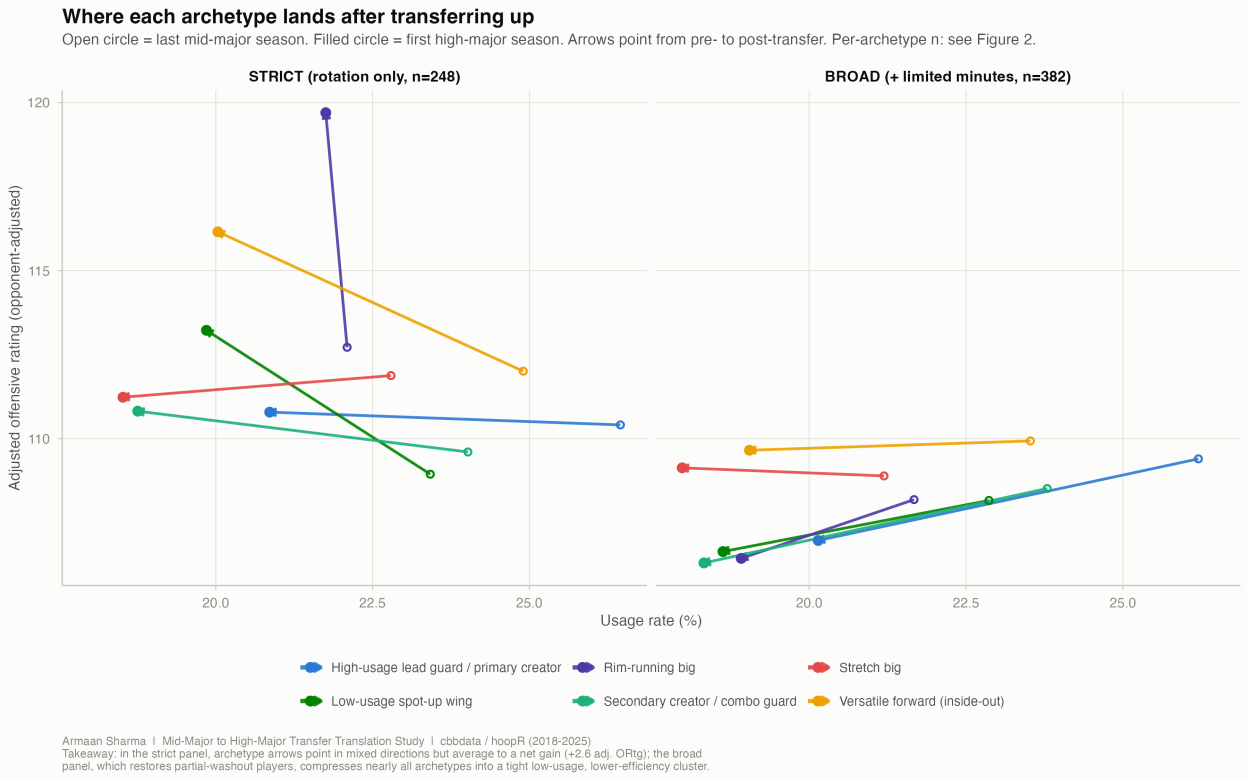


Figure 2: Where each archetype lands after transferring up, usage x adjusted-ORtg plane, strict vs broad.

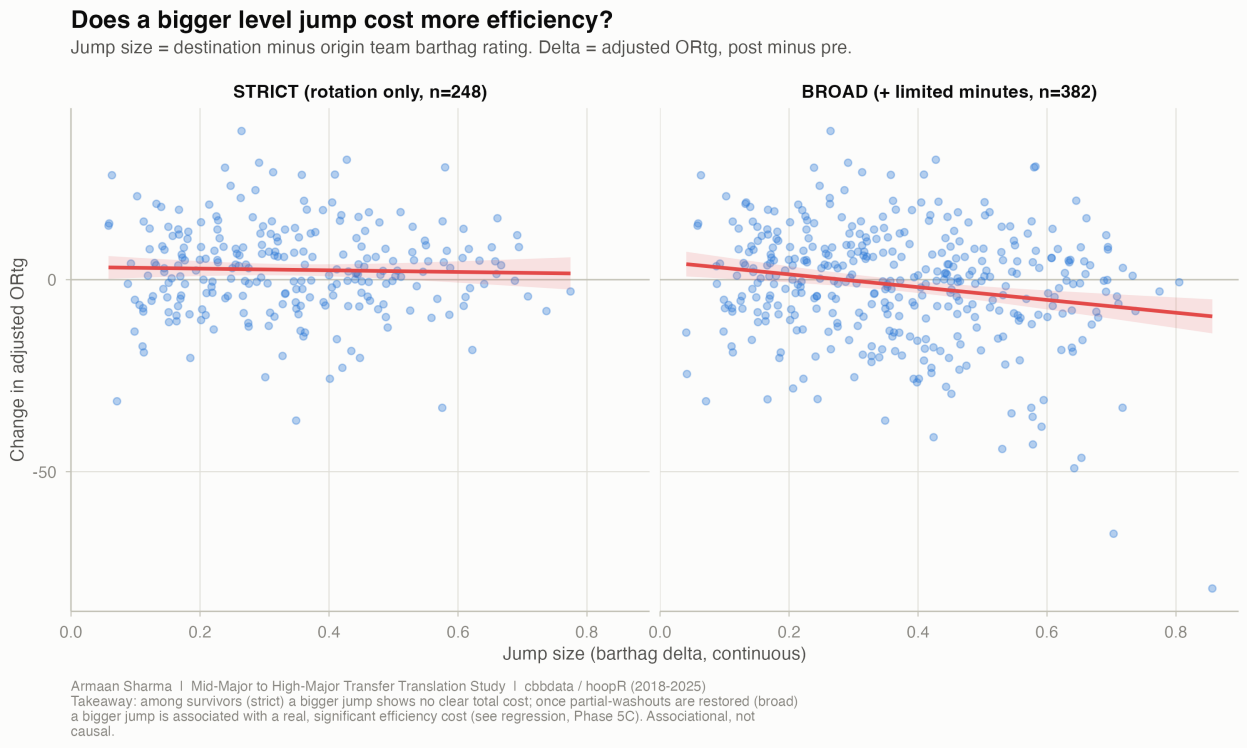


Figure 3: Change in adjusted ORtg vs. jump size, strict vs broad.

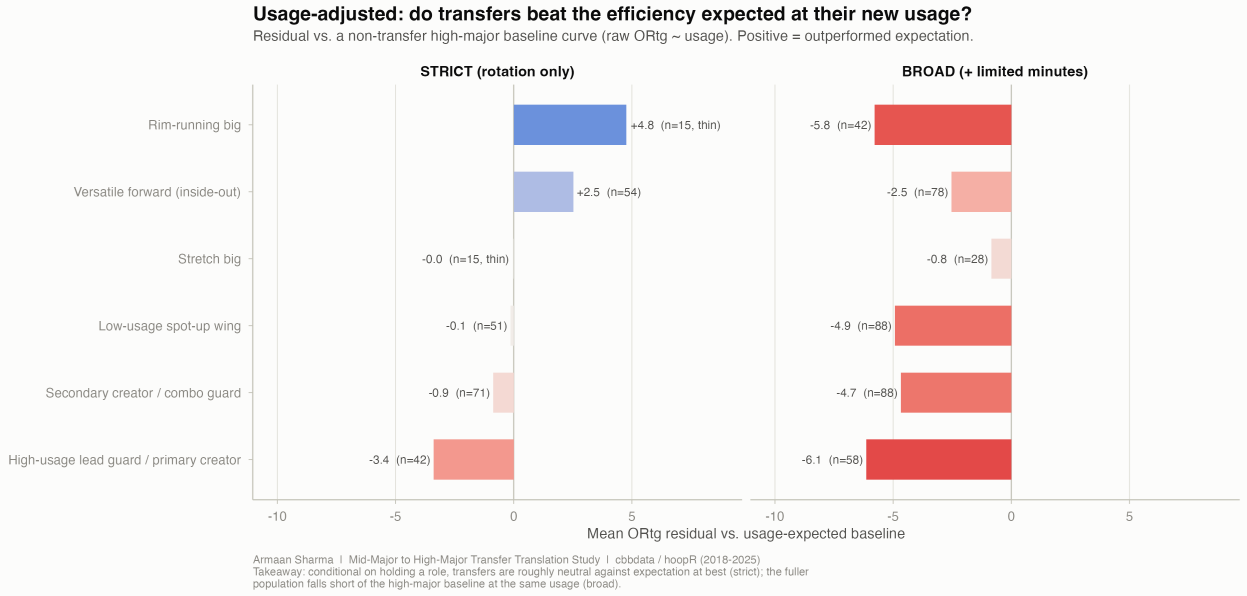


Figure 4: Usage-adjusted over/under-performance by archetype, strict vs broad.

real transfer makes, so coefficients are also expressed per an interquartile-range-sized jump.

Table 6: Jump-size coefficient across total- and direct-effect models, strict vs broad

| Sample | Effect | Coef. | SE | p | Sig. | Per IQR jump | n |
|---|---|---|---|---|---|---|---|
| Strict | Total | -2.03 | 4.10 | 0.621 | | -0.49 | 248 |
| Broad | Total | -21.74 | 4.37 | <.001 | *** | -5.27 | 382 |
| Strict | Direct | 9.14 | 4.04 | 0.025 | * | 2.22 | 248 |
| Broad | Direct | -13.66 | 4.71 | 0.004 | ** | -3.31 | 382 |

Total effect = jump_size only (headline). Direct effect = jump_size with usage_change held fixed (mechanism decomposition). Per IQR jump = coefficient x 0.2426 barthag (the interquartile range of jump size in this sample). Sig.: *** p<.001, ** p<.01, * p<.05.

Reading the primary (total effect) rows: among the players who held a role, making a typical (interquartile) further jump up the level ladder shows no detectable overall cost to adjusted ORtg. Once washouts are restored, that same typical jump is associated with roughly 5.3 fewer points of adjusted ORtg, a large and statistically clear effect. The direct-effect rows, which hold usage change fixed, show why the strict sample looks so different: conditional on retaining a similar share of usage, bigger jumpers among survivors actually perform better, a pattern consistent with selection (only unusually good players hold a role after an unusually big jump). Part of the broad sample's total cost runs through usage compression itself (mediation); part is a direct efficiency penalty beyond just losing touches.

# 4 Limitations

**Washout rate is a lower bound.** The transfer-detection method requires a player to appear in at least one hoopR/ESPN box score at the new school, including a zero-minute, dressed-but-did-not-play row, to be counted at all. A player who transfers and never appears in a single game (an immediate second transfer, a medical situation before the season starts, a player who leaves the program before ever dressing) is invisible to this pipeline entirely and is not represented anywhere in the 549-player cohort, not even as a labeled washout. True washout risk is at least 48.8% and is likely somewhat higher.

**Survivorship bias is the central methodological concern of this project, not a footnote.** The strict, rotation-only cut of the data is exactly the cut a casual read of box scores would produce, and it tells a materially different and more flattering story than the fuller picture. Every rate in this report is labeled strict or broad for this reason; a reader who drops that label from a headline number has recreated the bias this project exists to correct.

**The cohort is dominated by the post-waiver era.** The NCAA's one-time transfer waiver took effect for players whose first season at a new school was 2021-22 or later. 68.1% of this cohort's moves are post-waiver (n = 374); only 31.9% (n = 175) are pre-waiver, when most transfers had to sit out a year under the old rules. Pre- and post-waiver transfers are a structurally different population (waiver-eligible-and-played versus sit-a-year-then-play), and this report does not split the sample by era, so findings should be read as describing the pooled, waiver-era-dominated population rather than a stable rule regime.

**The 2020-21 COVID season adds noise to one slice of the window.** 170 of 549 cohort rows (31%) touch the 2020-21 season on either the pre- or post-transfer side, a season with an extra season of eligibility granted to all players and a compressed, disrupted schedule. This was not split out separately.

**Small archetype cells are flagged, not hidden, but should not be over-read.** Rim-running big and stretch big both have n = 15 in the strict sample, below the n = 25 threshold this project treats as a floor for interpretation. Their broad-sample equivalents (n = 42 and n = 28) clear the threshold and are more trustworthy; the strict-only readings for these two archetypes (in particular rim-running bigs appearing to beat usage-expectation in the strict cut) should be treated as suggestive at most.

**TS% and eFG% are not opponent-adjusted.** They are reported throughout as a descriptive cross-check alongside `adj_oe`, never as the primary cross-level efficiency comparison, because a mid-major player's shooting percentages partly reflect weaker defenses faced.

**Selection bias in who transfers up, and who plays, is not modeled.** Players who enter the portal and land a high-major offer are already a selected group (not a random sample of mid-major talent), and playing time itself reflects a coaching decision this analysis cannot separate from pure talent translation. No measure of scheme or role fit at the destination program is available.

**All findings are associational, not causal.** This is an observational cohort of players who selected into transferring; regression coefficients describe association after the stated controls, not the causal effect of forcing a player to change level.

## 5 Closing note

This report is intended as a quantitative supplement to scouting, to be read alongside film, medical, and fit evaluation, not in place of them. It quantifies a base rate and a discount; it does not evaluate any individual player's specific translation.