

NBA Forecasting and Market Pricing: A Chronological Evaluation of Pace-Efficiency Scores and Minutes-Based Player Distributions

Fourth & Value Research Group

fourthandvalue.com

October 18, 2026 · Version 1.0 · Model nba-v1.0

Technical report; not peer reviewed.

ABSTRACT

We evaluate Fourth & Value's first NBA forecasting system, nba-v1.0, on three seasons of public ESPN box scores (2023–24 through 2025–26; 3,958 games, 3,690 of them regular season, and 104,866 player rows). Hyperparameters are chosen on 2024–25 and the 2025–26 season is scored once. Game scores come from exponentially weighted pace and efficiency ratings, two ridge regressions and a bivariate-normal score matrix with empirical overtime, so moneyline, spread and total probabilities share one distribution. Player counts multiply projected minutes by per-minute rates and use a gamma–Poisson minutes mixture with a Gaussian copula for combined markets. On 2025–26 the moneyline log loss is 0.6025 against 0.6876 for a home-win-rate baseline. Rest and back-to-back terms add nothing measurable beyond pace and efficiency (joint log-score difference +0.0009 nats, interval $[-0.0059, 0.0089]$). Player forecasts beat a per-game average with no minutes model in every market, by 0.079 nats per appearance for points. Against 979 DraftKings lines that ESPN embedded in its game summaries, the model was worse than the no-vig price in all three game markets (moneyline +0.0367 nats, interval $[0.0232, 0.0511]$). Fitted on the first half of the season, the optimal blend put 5% weight on the model for moneylines and none for spreads and totals. Player-versus-opponent history, personal home/away splits, back-to-back splits and per-player bias corrections all failed to improve held-out player forecasts, repeating the result of FV-2026-03. No historical player-prop prices were available, so the player model has not been compared with a sportsbook. The system runs in shadow mode: forecasts are published, but NBA prices are not eligible for Top Picks.

Keywords: NBA; probabilistic forecasting; pace and efficiency; minutes projection; gamma–Poisson; Gaussian copula; chronological validation; market efficiency; player props.

1. Introduction: What the Findings Mean

Our NHL paper (FV-2026-02) separated two questions, and they apply to basketball too. How likely is an outcome, for example a player scoring more than 22.5 points? And does the sportsbook's price pay enough for that chance? A better answer to the first question does not produce a profitable answer to the second. NBA prices are among the most heavily bet in North America, they move on injury news within minutes, and they include the book's margin.

We built nba-v1.0 to answer the first question honestly and to measure how far it gets on the second. It forecasts the final score of every game and the points, rebounds, assists, threes, blocks, steals and turnovers of every listed player. It learns from earlier games only, fixes its settings on one season and is graded once on a later one. The results, in plain terms:

- **The game model knows basketball but not the news.** Team pace and scoring efficiency explain winners well: its moneyline forecasts are far better than always backing the home team. DraftKings' prices are better still, by a clear margin, in moneylines, spreads and totals. The model does not know who is injured or resting tonight, and the market does.
- **Minutes come first for players.** Forecasting how long a player will play, and then how much he produces per minute, beats averaging his recent box scores in every statistic. The gain is largest for points.
- **Old splits do not help.** As in hockey, football and baseball (FV-2026-03), a player's record against tonight's opponent, at home or away, or on the second night of a back-to-back did not improve the forecast once his form and minutes were known.
- **No edge is claimed.** We have no archive of historical player-prop prices, so the player model has not been tested against a book. Simple betting rules on the game model lost money at DraftKings' prices in 2025–26.

For readers, this means the NBA board shows forecasts as reference numbers. It does not show recommendations, and NBA prices do not enter Today's Picks. The sections below set out the method, every comparison we ran, and what remains unproven.

2. Data and Information Sets

2.1 Sources and coverage

All training data come from ESPN's public scoreboard and game-summary endpoints, which also serve our live pages [2]. A backfill script converts each final box score into team and player rows with stable ESPN identifiers and archives them as one compressed file per season [1]. The NBA Stats API was abandoned because it refuses requests from GitHub's hosted runners, where the daily job runs. ESPN files the in-season

tournament final as a regular-season game with a separate competition type; we label it `cup_final`. Players listed as inactive or who did not play keep a row with a reason code and no statistics. Four 2025–26 playoff summaries were unavailable and are missing; no regular-season game is missing.

Table 1. ESPN games and player rows by season. Missing: playoff summaries ESPN could not supply.

Season	Regular	Cup final	Play-in	Playoffs	Games	Player rows	Missing
2023–24	1,230	1	6	82	1,319	35,004	0
2024–25	1,230	1	6	84	1,321	35,198	0
2025–26	1,230	1	6	81	1,318	34,664	4

Team and player states are built only from regular-season and cup-final games. Playoff and play-in games are stored but excluded, for the reason given in Section 7.2. Unless stated otherwise, the test set is the 1,230 regular-season games of 2025–26.

2.2 Information available at forecast time

For a game on date d , every feature uses only games completed on earlier dates; same-day results never enter. Team and player states are updated after each date, so the held-out season is forecast sequentially with frozen regression coefficients and up-to-date ratings. ESPN does not publish revision times for box scores, so this is a *reconstructed* information set: it rules out using a game to predict itself, but it cannot certify that every historical number matches what was visible on the morning of the game.

The historical study does not know who was injured, resting or starting. Every player forecast is conditional on the player appearing, and the evaluation scores only players who did. Live inference partly closes this gap: it reads ESPN’s current rosters and injury report and withholds forecasts for players listed Out. Team forecasts make no injury adjustment, in history or live. That omission matters most in the market comparison of Section 8.

2.3 Market data

ESPN game summaries for 2025–26 carry a DraftKings moneyline, home spread and total with American prices for 979 of the 1,230 regular-season games [2]. Each game has one quote and its time is not stated; because summaries were captured after the game, it is most likely the last line before tip-off. Earlier seasons have no such field, and we hold no historical player-prop prices. The market comparison is therefore limited to one book, one season, three game markets and one quote per game.

3. Team Scoring and Coherent Settlement

3.1 Pace and efficiency ratings

Each team's state is an exponentially weighted sum over its earlier games, indexed by games played. With half-life h games, prior strength k games and league mean L , a rated quantity is:

$$w_j = 2^{-(n-j)/h}, \quad \tilde{x} = (\sum_j w_j x_j + k L) / (\sum_j w_j + k).$$

(1)

Possessions are estimated from the box score as FGA – OREB + TOV + 0.44·FTA and scaled to 48 minutes, so overtime does not inflate pace. Offensive and defensive efficiency are points scored and allowed per possession. At each new season the weights are multiplied by a carry factor c , which shrinks last season's ratings toward the league before any new game is played. Selection on 2024–25 chose $h = 10$ games, $c = 0.5$ and $k = 3$ (Table 2). For home team H and away team A , the unadjusted expected scores are:

$$\text{pace} = \text{pace}_H \text{pace}_A / \text{pace}_L, \\ \text{naive}_H = \text{pace} \cdot \text{off}_H \text{def}_A / \text{ppp}_L, \quad \text{naive}_A = \text{pace} \cdot \text{off}_A \text{def}_H / \text{ppp}_L.$$

(2)

3.2 Regression to regulation points

Two ridge regressions map the two naive scores, back-to-back indicators, days of rest (capped at four) and a neutral-site flag to each team's regulation points. Targets exclude overtime so that overtime can be modeled separately. Features are standardized on the training rows. The fitted coefficients appear in Table 5. The home intercept and the naive-score slopes near one say that the ratings carry most of the information; the back-to-back terms say a team on the second night scores about 1.4 fewer points and allows about 0.9–1.3 more.

3.3 One score matrix for every market

Regulation scores (H, A) are bivariate normal with the regression means and the residual covariance Σ , discretized to integers from 40 to 200. The fitted residual standard deviations are about 11.2 points per team, with correlation 0.21: fast, loose games raise both scores. Mass on the diagonal is a regulation tie. Each tie is passed through an empirical distribution of five-minute overtime scores from training games, up to four overtimes, with any remainder split by one point. The result is a single matrix $P(h, a)$ of final scores. Every game market reads from it:

$$\begin{aligned}\Pr(\text{home wins}) &= \sum_{h>a} P(h,a), \quad \Pr(\text{home covers } \ell) = \sum_{h-a+\ell>0} P(h,a), \\ \Pr(\text{over } L) &= \sum_{h+a>L} P(h,a), \quad \text{push mass} = \sum_{h-a+\ell=0} P \text{ or } \sum_{h+a=L} P.\end{aligned}$$

(3)

So the moneyline, every alternative spread and every alternative total are consistent with one another, and whole-number lines get an explicit push probability. Coherence is not realism. The normal model puts 2.48% of games in a regulation tie, while 4.39% of 2025–26 games went to overtime. End-of-game fouling and the incentive to tie rather than lose make close finishes more common than a smooth distribution suggests. The understated overtime mass slightly underprices overs and close spreads, and a later version should inflate the diagonal.

4. Player Minutes and Count Distributions

4.1 Minutes

Projected minutes $\hat{T}$ come from a ridge regression on the player's recent minutes (short and long weighted averages and his last game), his share of starts, the log of his games played, whether he recently changed teams, a long-layoff flag (more than ten days), a back-to-back flag and a blowout-risk term: the game model's expected absolute margin scaled by his usual minutes. Uncertainty around $\hat{T}$ is a normal distribution truncated to $[0.5, 48]$, with a standard deviation that depends on the projection: about 7 minutes for bench players and about 5 for 36-minute starters (Table 10). The positive back-to-back coefficient (+0.75 minutes) is a selection effect: rested stars do not appear in the data, and the teammates who do play absorb their minutes.

4.2 Per-minute rates, opponents and the recent-average blend

For each statistic s , a per-minute rate $\hat{r}_s$ is an exponentially weighted total divided by weighted minutes, with half-life 20 appearances and 100 minutes of a league prior for the player's position group. It is scaled by the expected game pace relative to his team's own pace, and multiplied by the opponent's allowed rate relative to the league, raised to an opponent weight ω ; selection chose $\omega = 1$, full adjustment (Table 3). A training-fitted linear blend then combines the structural mean $\hat{T} \hat{r}_s$ with the player's last-10 and last-30 per-game averages:

$$\mu_s = \beta_0 + \beta_1 \hat{T} \hat{r}_s (\text{pace}/\text{pace}_{\text{own}}) (\text{opp}_s/\text{lg}_s)^\omega + \beta_2 \bar{y}_{s,10} + \beta_3 \bar{y}_{s,30}.$$

(4)

The structural term carries most of the weight (β_1 between 0.83 and 1.03 across statistics); the recent averages correct the structural mean slightly. The blended mean divided by $\hat{T}$ becomes the final per-minute rate.

4.3 Counts, dispersion and combined markets

Given minutes T drawn from the minutes distribution, a count is negative binomial with mean Tr_s and gamma shape κ_s , so $\text{Var}(Y | T) = \mu + \mu^2/\kappa_s$ [3]. The shapes are fitted by maximizing the likelihood of the whole predictive mixture over minutes, which keeps the minutes uncertainty from being counted twice. Points need strong extra dispersion ($\kappa = 5.4$); turnovers are effectively Poisson ($\kappa \approx 1,982$).

Points + rebounds + assists and the other combined markets need a joint distribution. A simulation draws one minutes value per player, then correlated gamma multipliers through a Gaussian copula whose correlation matrix is estimated from training residuals, then Poisson counts. Every statistic in a draw shares the same minutes. Finally, each market's probability passes through a logit slope fitted on 2024–25 pseudo-lines:

$$p' = \sigma(b_m \cdot \text{logit } p).$$

(5)

A slope b below one pulls probabilities toward 50%. The fitted slopes are 0.78 for points and 0.75–0.83 for the three combined markets containing points, which says the raw tails were too confident. The other markets have slopes of 1.00–1.08.

5. Markets, Pushes and Price Evaluation

For a two-way market with American prices converted to implied probabilities q_1, q_2 , the no-vig market probability is $q_1/(q_1+q_2)$. On a whole-number line this is the probability of winning given no push. The model's comparable probability is its win mass divided by one minus its push mass. With win, push and loss probabilities w, u, l , a one-unit bet at decimal odds d is worth:

$$\text{EV}(d) = w(d - 1) - l = wd - (1 - u), \quad d_{\text{fair}} = (1 - u)/w.$$

(6)

The live board computes this value for every quote but does not display it, because the model has not been graded against prices (Section 10). In this paper, pushes are excluded when comparing the model with the market, and refunded in the policy simulation.

6. Evaluation Design and Statistical Estimands

6.1 Chronological protocol

The protocol is fixed in the training script [1]. Selection fits on 2023–24 from December 1, when team ratings have settled, and scores candidates on 2024–25. The test fits the selected configuration on both earlier seasons and scores 2025–26 once. The production artifact is then refit on all three seasons. The grids searched were team half-life {10, 20, 40}, carry {0.25, 0.5} and prior games {3, 8}; player rate half-life {20, 50}, prior minutes {100, 300} and opponent weight {0, 0.5, 1}. The game model’s feature set was fixed in advance, not searched. The calibration slopes and blend coefficients were fitted on the selection data only.

Table 2. Team-state selection: 2024–25 joint-score NLL per game (nats), fit on 2023–24. Lower is better; first row selected.

Half-life (games)	Season carry	Prior games	Validation NLL
10	0.5	3	7.6561
10	0.25	3	7.6567
20	0.25	3	7.6580
10	0.5	8	7.6606
10	0.25	8	7.6609
20	0.25	8	7.6610
20	0.5	3	7.6619
20	0.5	8	7.6646
40	0.25	3	7.6689
40	0.25	8	7.6701
40	0.5	3	7.6792
40	0.5	8	7.6796

Table 3. Player-rate selection: 2024–25 count NLL summed over seven statistics. First row selected.

Rate half-life (apps)	Prior minutes	Opponent weight w	Validation NLL sum
20	100	1.0	11.6329
20	100	0.5	11.6338
50	100	1.0	11.6341
50	100	0.5	11.6368
50	300	1.0	11.6434
20	100	0.0	11.6443
50	100	0.0	11.6465
50	300	0.5	11.6471
20	300	1.0	11.6559
50	300	0.0	11.6561
20	300	0.5	11.6592
20	300	0.0	11.6680

Differences inside each grid are small: the twelve team configurations span 0.024 nats per game, and the best player configuration beats the runner-up by less than 0.001 nats summed over seven statistics. The selected settings are reasonable choices, not sharply identified ones.

6.2 Scores and uncertainty

Score distributions are graded by negative log likelihood (NLL) of the observed final score or count; binary events by log loss and Brier score, which are proper scoring rules [4]. All comparisons are paired on identical games or appearances. Intervals are 95% percentile intervals from 500 bootstrap replicates that resample whole games, keeping every player and market from one game together. They do not reflect the choice of specification, serial correlation between games of the same team, or simultaneous coverage across the many comparisons below.

Player comparisons use the 23,979 test appearances whose player had at least 20 earlier appearances, so that the history-based baselines are defined. Because no historical prop lines exist, binary player results use a *pseudo-line*: the floor of the player’s last-10 average plus one half. Pseudo-lines resemble real prop lines in that they sit near recent form, but they are not book prices. Combined-market results use a seeded sample of 3,000 appearances because each requires simulation.

7. Forecast Results and Ablations

7.1 Game forecasts

Table 4. 2025–26 game forecasts, 1,230 games. Bias is actual minus expected total points; early is the first three weeks (156 games).

Model	Joint NLL	ML log loss	ML Brier	Total RMSE	Margin RMSE	Total bias	Early bias
Home-win rate	–	0.6876	0.2472	–	–	–	–
Pace × efficiency only	7.7459	0.6011	0.2067	18.97	14.59	+0.97	+2.88
v1 (adds rest, back-to-back)	7.7469	0.6025	0.2076	18.97	14.59	+0.96	+2.81
v1 with playoff games in states	–	–	–	19.02	–	+1.08	+4.43

The moneyline improvement over the home-win-rate baseline is large and precise: -0.0851 nats per game, interval $[-0.1066, -0.0650]$. The rest and back-to-back terms do not help on the test season. A model with only the two naive scores and the neutral flag has a slightly *better* joint log score (difference $+0.0009$ nats, interval $[-0.0059, 0.0089]$) and moneyline log loss. The rest effects in Table 5 are real in training, but rest is also correlated with schedule density that pace ratings already absorb, and the gain does not survive out of sample. We report this rather than drop the terms after seeing the test season; v1 keeps its preselected feature set.

The regression adjustment barely changes totals: RMSE 18.97 against 19.01 for the unadjusted pace × efficiency product. Total errors are roughly 19 points per game whichever way the mean is built; most of a total’s uncertainty is not predictable from team ratings.

Table 5. v1 regulation-point regressions (fit 2023–24 and 2024–25), per unit of each feature.

Feature	Home points	Away points
Intercept (points)	114.294	112.558
Naive home score	0.955	-0.058
Naive away score	0.010	0.944
Home on back-to-back	-1.377	0.857
Away on back-to-back	1.303	-1.422
Home days of rest	0.130	-0.385
Away days of rest	0.334	0.743
Neutral site	-9.049	4.256

Table 6. 2025–26 moneyline reliability: home-win forecasts by bin.

Forecast bin	Games	Mean forecast	Home won
0.1–0.2	38	0.161	0.210
0.2–0.3	95	0.259	0.179
0.3–0.4	142	0.355	0.366
0.4–0.5	204	0.452	0.456
0.5–0.6	248	0.551	0.536
0.6–0.7	211	0.646	0.678
0.7–0.8	161	0.747	0.789
0.8–0.9	112	0.845	0.830

7.2 Early-season totals and playoff contamination

Games in the first three weeks finished 2.81 points above the model’s expected total, against +0.96 over the whole season. The carry factor pulls ratings toward the league mean after the summer, while league scoring has drifted upward from season to season. A first version also built team states from playoff games. Playoff games are 12–17 points lower scoring and come at the end of the season, so they dragged opening-night ratings down. With playoff games in the states, early-season bias was +4.43 points; removing them gave +2.81. The low-scoring study (FV-2026-04) found that the closing market already prices playoff scoring [6]; this result is the modeling counterpart. Playoff basketball is a different scoring environment and should not inform regular-season ratings without an explicit adjustment.

7.3 Player count forecasts

Table 7. 2025–26 count NLL per appearance (nats), 23,979 appearances. Lower is better.

Statistic	Recency 30	Direct per-game	Minutes × rate, Poisson	Structural (no blend)	v1
Points	3.2395	3.1967	3.2731	3.1226	3.1179
Rebounds	2.2542	2.2278	2.1803	2.1780	2.1782
Assists	1.8342	1.8219	1.7893	1.7919	1.7872
Threes	1.3585	1.3498	1.3396	1.3405	1.3375
Blocks	0.8348	0.8273	0.8212	0.8200	0.8201
Steals	1.1761	1.1677	1.1570	1.1566	1.1567
Turnovers	1.4183	1.4121	1.3900	1.3907	1.3900

Each column removes or replaces one part of v1. *Recency 30* is the player’s last-30 per-game average with v1’s dispersion. *Direct per-game* smooths per-game counts with the same half-life and fits its own negative-binomial shape, but has no minutes model. *Minutes × rate, Poisson* keeps the minutes mixture but drops the extra dispersion. *Structural* is v1 without the recent-average blend.

Table 8. Paired differences in count NLL, $v1$ minus each alternative, with 95% game-bootstrap intervals. Negative favors $v1$.

Statistic	vs recency 30	vs direct per-game	vs Poisson	Blend effect
Points	-0.1216 [-0.1302, -0.1129]	-0.0788 [-0.0845, -0.0744]	-0.1553 [-0.1679, -0.1437]	-0.0047 [-0.0061, -0.0035]
Rebounds	-0.0759 [-0.0825, -0.0695]	-0.0496 [-0.0537, -0.0449]	-0.0021 [-0.0032, -0.0009]	+0.0003 [-0.0003, 0.0008]
Assists	-0.0470 [-0.0516, -0.0424]	-0.0347 [-0.0382, -0.0312]	-0.0021 [-0.0032, -0.0011]	-0.0047 [-0.0059, -0.0036]
Threes	-0.0210 [-0.0244, -0.0181]	-0.0123 [-0.0149, -0.0100]	-0.0021 [-0.0030, -0.0013]	-0.0030 [-0.0036, -0.0025]

Three findings. First, separating minutes from per-minute production beats the direct per-game model in every statistic, by 0.079 nats per appearance for points and 0.012 for threes; every interval excludes zero. This repeats the NHL result, where time on ice gave the largest player gain. Second, extra dispersion matters for points and little else: the Poisson version is 0.155 nats worse for points but within 0.002 for the other statistics. Points come in bunches of two and three, so a scorer's total varies far more than a Poisson count would. Third, the recent-average blend gives a small, consistent gain for points, assists and threes, about 0.003–0.005 nats, and none for rebounds.

Table 9. Distribution diagnostics, 2025–26. Coverage of the central 80% discrete interval; κ is the gamma shape.

Statistic	Observed zero	Predicted zero	80% coverage	Shape κ
Points	9.16%	2.65%	82.8%	5.4
Rebounds	9.34%	7.79%	88.8%	28.5
Assists	21.92%	20.42%	91.3%	26.3
Threes	40.87%	39.00%	93.7%	14.6
Blocks	67.81%	67.05%	95.5%	8.5
Steals	49.07%	48.38%	94.7%	18.0
Turnovers	34.60%	34.39%	93.7%	1,981.6

Zero rates match closely except for points: 9.16% of appearances ended scoreless against 2.65% predicted. Most scoreless games come from brief garbage-time or injury-shortened appearances, which the minutes distribution understates. These are rarely the players with posted props. The nominal 80% central intervals cover 83–95% of outcomes; integer intervals on low counts are necessarily conservative, so high coverage for blocks or steals is not evidence of a poorly sized distribution.

Table 10. Minutes projection, 2025–26 appearances, against the player’s last-five-game average.

<i>Measure</i>	<i>v1 regression</i>	<i>Last 5 average</i>
Mean absolute error (min)	4.86	4.96
RMSE (min)	6.29	6.46
Residual SD at 2 projected min	6.88	—
Residual SD at 10 projected min	7.36	—
Residual SD at 18 projected min	7.00	—
Residual SD at 26 projected min	6.18	—
Residual SD at 34 projected min	5.29	—
Residual SD at 38 projected min	5.25	—

The minutes regression improves on a five-game average by about 0.1 minutes of absolute error. That is a small gain because most minute changes are driven by events the model cannot see, such as foul trouble, blowouts, injuries and rotation decisions. Minutes error is the largest single source of player-forecast uncertainty.

7.4 Over/under probabilities at pseudo-lines

Table 11. Over/under log loss at pseudo-lines (floor of last-10 average + 0.5), 2025–26. Slopes fitted on 2024–25.

<i>Market</i>	<i>Appearances</i>	<i>Over rate</i>	<i>Baseline</i>	<i>v1 raw</i>	<i>v1 calibrated</i>	<i>Slope</i>
Points	23,979	45.7%	0.6957	0.6749	0.6741	0.777
Rebounds	23,979	44.3%	0.6819	0.6608	0.6607	1.020
Assists	23,979	42.8%	0.6789	0.6573	0.6575	1.069
Threes	23,979	36.5%	0.6067	0.5961	0.5957	1.036
Blocks	23,979	27.9%	0.5592	0.5478	0.5479	1.078
Steals	23,979	38.7%	0.6447	0.6297	0.6296	1.018
Turnovers	23,979	41.9%	0.6629	0.6423	0.6426	1.075
Pts + Reb + Ast	3,000	47.4%	0.7038	0.6759	0.6751	0.830
Pts + Reb	3,000	46.6%	0.7082	0.6792	0.6777	0.753
Pts + Ast	3,000	46.1%	0.7026	0.6776	0.6764	0.759
Reb + Ast	3,000	46.4%	0.6970	0.6729	0.6730	1.008

Against the same pseudo-lines, the calibrated model beats an empirical baseline, the share of the player’s last 30 games (within a year) over the line with a one-game Laplace correction, in all eleven markets. For points the margin is 0.6741 against 0.6957. The calibration slopes were fitted on 2024–25 and applied unchanged to 2025–26. Here the model is graded on lines that follow recent form; a book’s line also knows the opponent, the injury report and the betting flow.

Table 12. Combined markets: joint simulation (shared minutes, copula) against independent marginals, 3,000 sampled appearances.

Market	Joint log loss	Independent	Joint – independent [95%]	Observed var.	Joint var.	Independent var.
Pts + Reb + Ast	0.6820	0.6853	–0.0033 [–0.0050, –0.0016]	60.6	73.3	60.7
Pts + Reb	0.6805	0.6824	–0.0019 [–0.0031, –0.0007]	51.4	64.8	57.2
Pts + Ast	0.6795	0.6799	–0.0004 [–0.0012, 0.0004]	43.7	57.1	53.9
Reb + Ast	0.6711	0.6736	–0.0025 [–0.0039, –0.0012]	11.5	12.0	10.2

Joint simulation beats independent marginals for points + rebounds + assists (–0.0033 nats, interval [–0.0050, –0.0016]), points + rebounds and rebounds + assists. The difference is not resolved for points + assists. Shared minutes make a player’s statistics rise and fall together. The joint model nonetheless overstates combined variance (73.3 predicted against 60.6 observed for P+R+A), and the calibration slope of 0.83 partly absorbs this. The slopes were fitted to the joint simulation, which favors it in this comparison. Within one player, the joint model is the coherent choice; no dependence across different players is modeled, so same-game parlays are unsupported.

8. Comparison with DraftKings Prices

8.1 Model against market

For each of the 979 games with a DraftKings line, we compare the test-season model with the no-vig market probability on the same binary event. Games that pushed are excluded from spread and total comparisons.

Table 13. v1 against DraftKings no-vig probabilities, 2025–26, pushes excluded. Blend weight fitted on the first half; second-half market / blend log loss.

Market	Games	Model	Market	Model – market [95%]	Mean gap	Blend weight	2nd half
Moneyline	973	0.6077	0.5710	+0.0367 [0.0232, 0.0511]	8.6%	0.05	0.5001 / 0.5010
Spread	978	0.7268	0.6930	+0.0338 [0.0191, 0.0482]	8.9%	0.00	0.6933 / 0.6933
Total	978	0.7174	0.6929	+0.0245 [0.0124, 0.0368]	7.6%	0.00	0.6927 / 0.6927

The market is better in all three markets, by 0.024–0.037 nats per game, and every interval excludes zero. On spreads and totals the market’s log loss is close to that of a coin flip (0.693). The model’s is worse than a coin flip: it disagrees with the line by about 9 percentage points on average, and those disagreements carry no information. In its most confident spread bin (forecast 0.73) the home side covered 31% of the time; in its least confident (0.27) it covered 51%. That is the pattern of a model missing information that the line already contains, chiefly who is playing.

The blend test asks whether the model adds anything to the market. A logit blend weight was fitted on the first half of the season and scored on the second. It put 5% weight on the model for moneylines and none for spreads and totals. On the second half the moneyline blend scored 0.50103 against the market’s 0.50006, so

even that 5% did not help. There is no evidence that v1 contains information the closing line lacks.

8.2 An exploratory betting rule

Table 14. Exploratory flat-stake rule at DraftKings prices, 2025–26: bet sides whose model probability beats break-even by the threshold. Pushes refunded.

Market	Threshold	Bets	ROI	ROI 95% interval
Moneyline	0.0 pp	828	−12.9%	[−23.3%, −3.9%]
Moneyline	2.5 pp	662	−16.2%	[−26.4%, −6.0%]
Moneyline	5.0 pp	497	−16.3%	[−28.3%, −3.0%]
Spread	0.0 pp	832	−8.6%	[−14.8%, −2.1%]
Spread	2.5 pp	659	−7.8%	[−14.7%, −0.4%]
Spread	5.0 pp	502	−9.8%	[−18.7%, −0.8%]
Total	0.0 pp	803	−5.1%	[−11.1%, 1.4%]
Total	2.5 pp	608	−4.2%	[−11.5%, 2.9%]
Total	5.0 pp	427	−8.4%	[−17.8%, −0.2%]

Betting one unit on every side whose model probability exceeded the vigged break-even by a threshold lost money in every market and at every threshold. These rules were defined after the price data were seen, so they are illustrations, not a pre-registered test. The losses are what Section 8.1 predicts: disagreement with a sharp line, without new information, buys the vig and adverse selection. Totals lost least, and their intervals include break-even at the two lower thresholds. Moneyline rules lost most, since disagreements tend to favor underdogs at long prices.

9. Matchups, Venue, Back-to-Backs and Player Bias

FV-2026-03 asked whether a player’s history against one opponent, or his own home/away split, predicts his next game in the NFL, NHL and MLB [5]. We applied the same method to NBA points, rebounds, assists and threes, and added back-to-back splits. For a player with earlier actual total A and forecast total E within a split, the adjusted forecast multiplies v1’s mean by a shrunk ratio:

$$r_{\text{split}} = (A_{\text{split}} + k \cdot r_{\text{all}}) / (E_{\text{split}} + k), \quad r_{\text{all}} = (A + k) / (E + k).$$

(7)

The overall ratio r_{all} is a per-player bias correction; split ratios shrink toward it. The base forecast is v1 fitted on 2023–24. The shrinkage k is chosen from $\{1, 3, 10, 30, 100, 300, \infty\}$ on 2024–25, where ∞ means no adjustment, and the result is scored on 2025–26. Only earlier dates enter A and E .

Table 15. Shrunk player histories over v1, scored on 2025–26. Change in count NLL against no split; positive is worse. “not used” is $k = \infty$.

Statistic: adjustment	Selected k	Test change [95%]	Best finite k	Change at that k [95%]
Points: player bias	not used	no change	—	—
Points: home/away split	not used	no change	300	+0.00146 [0.00000, 0.00280]
Points: vs this opponent	not used	no change	300	+0.00083 [0.00009, 0.00157]
Points: back-to-back split	not used	no change	300	+0.00085 [−0.00057, 0.00222]
Rebounds: player bias	not used	no change	—	—
Rebounds: home/away split	not used	no change	300	+0.00111 [0.00031, 0.00200]
Rebounds: vs this opponent	300	+0.00008 [−0.00030, 0.00052]	300	+0.00008 [−0.00032, 0.00052]
Rebounds: back-to-back split	not used	no change	300	+0.00035 [−0.00058, 0.00122]
Assists: player bias	not used	no change	—	—
Assists: home/away split	not used	no change	300	+0.00002 [−0.00058, 0.00058]
Assists: vs this opponent	not used	no change	300	+0.00000 [−0.00024, 0.00025]
Assists: back-to-back split	not used	no change	300	+0.00028 [−0.00035, 0.00086]
Threes: player bias	not used	no change	—	—
Threes: home/away split	not used	no change	300	+0.00016 [−0.00017, 0.00049]
Threes: vs this opponent	300	+0.00003 [−0.00009, 0.00016]	300	+0.00003 [−0.00008, 0.00017]
Threes: back-to-back split	not used	no change	300	+0.00017 [−0.00016, 0.00051]

Selection set $k = \infty$, no adjustment, for every statistic and split except opponent history for rebounds and threes, where it chose the weakest finite setting. Neither changed the test score measurably. To check that infinite shrinkage was not hiding a real effect, the right-hand column scores the best finite k anyway: every split made test forecasts slightly worse or no better. Players who had beaten their forecast against an opponent, at home, or on back-to-backs did not keep doing so.

Two results differ from the earlier study. First, the per-player bias correction that helped hockey and became nhl-v2.4 does not help here. The recent-average blend in (4) already moves each player’s forecast toward his own recent results. Second, there is no league-wide back-to-back error left to correct: in 2025–26, players on back-to-backs reached 100.1% of their forecast points and rested players 99.1%. The team model includes back-to-back terms and the minutes model includes a back-to-back flag, and FV-2026-04 found that closing totals already reflect back-to-backs [6].

10. Shadow-Mode Operation

Since October 17, 2026, the daily NBA job refreshes ESPN history, fits nothing, loads the frozen production artifact and adds v1 probabilities to every priced quote. Game forecasts carry the score-matrix probabilities and push masses; player forecasts use ESPN’s current roster and injury report. Players listed Out are with-

held, as are players who cannot be matched unambiguously to a current roster and players with too few NBA appearances. Each withheld quote records its reason. The cards show the v1 line next to the market price; expected value is computed but not displayed.

Shadow mode is deliberate. The feed never marks NBA forecasts as model-checked, so the editorial writer and Today's Picks skip NBA. The [current daily process](#) describes the eligible sports. Grading v1 against the 2026–27 prices it actually sees, with the timestamps and injury information of each quote, is the evidence needed before NBA becomes eligible. This paper's results set the bar: the game model would have to show that it adds information to the closing line, which it did not do in 2025–26.

11. Threats to Validity and Remaining Research

Availability and lineups. The historical study knows nothing about injuries, rest or starting lineups, and team forecasts ignore them even live. That is the most likely reason the market wins, and adding availability is the most important next step. A player-level team model, built by adding up projected minutes and per-minute contributions of the players expected to play, would let injury news move team forecasts.

Market data. One book, one season, one quote per game of unstated time, and no player props. The model–market differences are precise for this sample, but they cannot speak to opening lines, other books, or props, where markets are thinner and more often stale. The live system now archives every quote it prices, which will support a dated comparison for 2026–27.

Participation. Player results are conditional on appearing. A prop is usually voided if the player does not play, but a player who leaves early through injury or ejection stays in the evaluation. The minutes distribution only partly captures that risk.

Distributional shape. The normal score model underprices overtime; the joint player simulation overstates combined variance, and calibration slopes correct both only partly. Pseudo-lines measure calibration near recent form, not at the lines books post.

Statistical precision and multiplicity. Hundreds of comparisons appear in this report. Intervals are per comparison and conditional on the selected specification; small differences near zero, such as the rest-term ablation or individual matchup splits, should be read as “no measurable effect,” not as estimated harms. Three seasons of ESPN data are a short history; seasons before 2023–24 would strengthen selection and the test.

Next evidence required. Prospective, timestamped grading of v1 against the quotes it sees in 2026–27; injury-aware team forecasts tested on held-out games; a fix for overtime mass; and, if obtainable, a historical player-prop archive. No recommendation threshold should be enabled because the system is running.

12. Conclusion for Readers and Analysts

nba-v1.0 is a sound basketball model with a clear blind spot. It turns team pace and efficiency into a coherent distribution of final scores, and it forecasts player statistics better than recent averages by projecting minutes first. These are useful reference numbers for the board.

It does not beat DraftKings. The market knew more than the model about every kind of game line in 2025–26, and simple rules that bet the model’s disagreements lost money. The most likely reason is that the market knows who is playing tonight and the model’s team forecasts do not. Head-to-head history, home/away splits and back-to-back splits do not close the gap.

Use the NBA forecasts to see where a price sits against a transparent baseline, not as a signal to bet. NBA prices will become eligible for picks only after the model has been graded against real 2026–27 prices and shows that it adds something the market lacks.

Appendix A. Reproducibility and Artifact Lineage

This report describes model `nba-v1.0`, whose implementation was merged at commit `ded4221d`. The study script `scripts/research/nba/study.py` writes `reports/nba-paper/evidence.json`, and every table here is generated from that file. The study adds one change to the model code: `fit_game` accepts a feature subset for the ablation in Table 4, and production output is unchanged. The downloadable evidence JSON contains all diagnostics and the SHA-256 hashes of the three season archives, the production model and the v1 evaluation.

The environment pins Python 3.11, NumPy 1.26.4 and SciPy 1.17.1. The season archives under `artifacts/nba/history/` are committed. The market section needs the cached ESPN summaries, which are not committed; the backfill command restores them from ESPN without cost. From a checkout:

```
python3.11 -m venv .venv
.venv/bin/pip install -r requirements-nhl.txt
.venv/bin/python scripts/nba/espn_history.py --season 2025-26
.venv/bin/python scripts/nba/train.py
.venv/bin/python scripts/research/nba/study.py
python3 scripts/research/build_nba_paper.py
node scripts/research/render_nba_paper.cjs
```

The study runs on a laptop CPU in about an hour, most of it spent on combined-market simulation and matchup histories. Random seeds are fixed. ESPN can revise historical box scores, so a fresh download is a new data vintage; the hashes in the evidence JSON identify the one used here.

References and Research Materials

- [1] Fourth & Value Research Group (2026). NBA v1 implementation, model card and evaluation.
- [2] ESPN. NBA scoreboard and game-summary endpoints (site.api.espn.com). Public, undocumented JSON; retrieved October 2026. Box scores and the DraftKings lines in the summary `pickcenter` field.
- [3] SciPy developers. Negative-binomial distribution reference. Accessed October 18, 2026; computation pinned to SciPy 1.17.1.
- [4] Gneiting, T., and Raftery, A. E. (2007). Strictly Proper Scoring Rules, Prediction, and Estimation. *Journal of the American Statistical Association*, 102(477), 359–378. DOI: 10.1198/0162145060000001437.
- [5] Fourth & Value Research Group (2026). Player Matchups and Home Field. Technical Report FV-2026-03.
- [6] Fourth & Value Research Group (2026). What Makes Games Go Under? Technical Report FV-2026-04.
- [7] Fourth & Value Research Group (2026). NHL Forecasting and Market Pricing. Technical Report FV-2026-02.

Research status: shadow-mode forecasts for NBA game lines and seven player statistics with four combined markets. No betting advantage over the market is established, and no player-prop price comparison has been made. This report is an internal technical publication, not external peer review.