Cross-Sport Shot Analysis

Kanaan Law, Kaleb Coleman, and Drianna Walker
Department of Mathematics and Statistics

Northern Arizona University
Office of Undergraduate Research and Creative Activity

Northern Arizona University
Department of Mathematics and Statistics

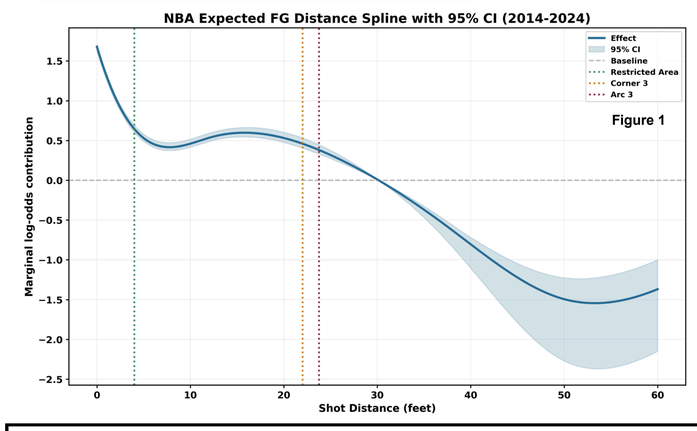
Abstract

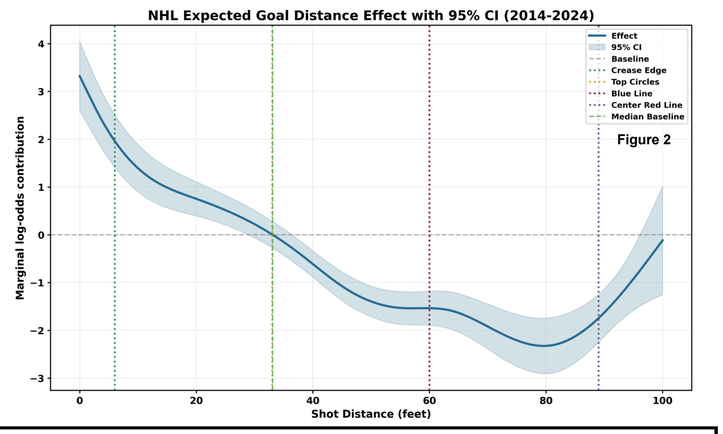
This project develops parallel expected scoring models for the NBA and NHL to evaluate shooting efficiency across two major professional sports. Using NBA optical-tracking data and NHL MoneyPuck shot data, we estimate shot-by-shot scoring probability (xFG and xG). We then quantify performance using residual metrics such as Points/Goals Over Expected (POE) and a Shot Difficulty Index (SDI). By standardizing shot context, difficulty, and efficiency across leagues, we identify players who consistently over- or under-perform expectations and highlight structural differences in shot environments.

Goals and Difficulties

Modern sports analytics increasingly rely on expected-value modeling to distinguish shot quality from shot results, since raw percentages often obscure how difficult a player's attempts truly are and thus a player's true value can get lost. Comparing the NBA and NHL requires navigating two radically different scoring landscapes; one with only a few makes per game, the other with well over a hundred; Along with mismatched definitions of assists, divergent efficiency metrics, and highly specialized player roles. Even with these structural differences, aligning the models and examining their residual patterns revealed genuine cross-sport parallels in how players create value from difficult attempts. By standardizing the model outputs, we established a shared framework that makes NBA and NHL shots directly comparable, allowing us to analyze shot difficulty and scoring efficiency metrics across both sports despite their fundamentally different contexts.

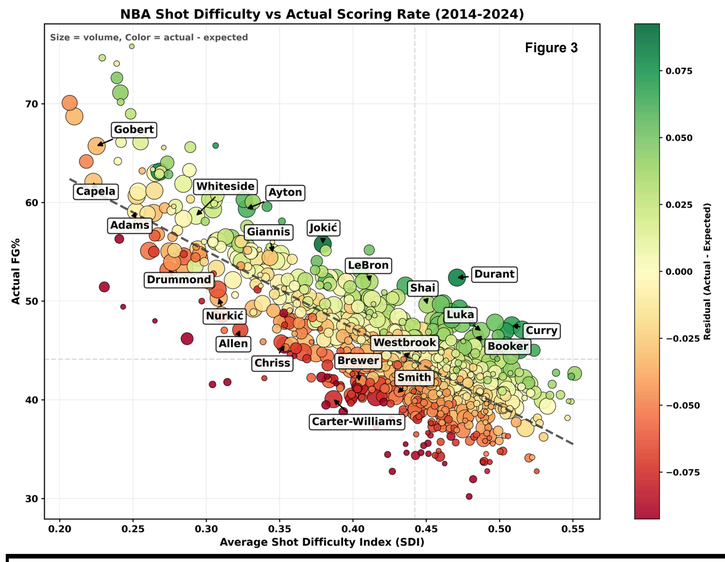
GAM for Shot Distance

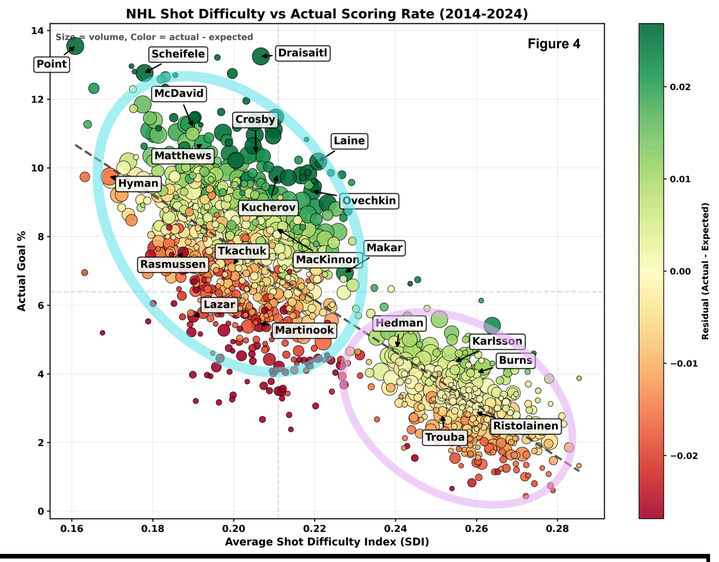




The generalized additive model (GAM) illustrates the marginal effect of shot distance on goal probability, controlling for contributing factors. The blue line represents the estimated smooth function and the shaded ribbon represents the 95% confidence interval. Dashed lines representing significant court/rink markings are also included for refrence.

Shot Difficulty Index





The scatterplots above represent the Shot Difficulty Index on the x-axis (higher = harder shots) versus actual goal percentage on the y-axis. Each point represents a player, with the bigger the point size the bigger the shot volume. The color represents residual goal efficiency (Actual - Expected), the green points are overperformers and the red are underperformers. The dotted line is a simple regression line. Key patterns: Lower-right green players thrive on difficult shots, lower-right red players attempt tough shots but convert poorly, left-side red players struggle on easier attempts. The circles on the NHL plot represent position clusters that formed.

Data Snapshot

	NBA	NHL	Description/Importance
Spatial Location	LOC_X, LOC_Y	xCord, yCord	Where the player shot the ball/puck
Shot Distance	shot_distance_feet	shotDistance	Distance the player was from the hoop/net when play was made
Shot Angle	shot_angle	shotAngle	Angle the ball/puck was shot at in relation to the goal (straight on = 0°)
Game Timing	seconds_in_period, PERIOD, game_minutes_elapsed	period, time	How far into the game (min and sec for NBA, sec for NHL) the play happened
Shot Type	shot_type_cat plus dunk/layup/hook/floater/jump indicators	shotType, wrist/snap/slap/backhand indicators	Helped determine player shot specializations, game pace, varying strategies
Extra Context	SHOT_ZONE_BASIC, is_clutch	shotRebound, shotGoalieFroze, shotRush	Features incorporated in our SDI equation, varying metrics which had impacts on goal scoring

Results

Figure 1 (NBA GAM):

- Attempts just inside 10 feet and at the 3-point line yield similar efficiency.
- Further analysis revealed that beyond 40 feet, only elite shooters like Curry and Durant take shots, while the confidence interval reflects average players who rarely attempt from that range.

Figure 2 (NHL GAM):

- Median shot distance, "good shot" threshold, and top of the faceoff circles align closely, reinforcing model consistency.
- Slight efficiency bump after the defensive zone: likely due to empty-net and defensive player goals.

Figure 3 (NBA SDI):

- No clusters: shot difficulty and efficiency scale continuously. This shows the NBA is much less specialist oriented and position versatility is more prioritized.
- Players above the regression line can still be inefficient if they miss easy shots.
- Elite scorers (e.g., Kevin Durant, Nikola Jokić) appear in expected positions: high difficulty, high efficiency while High-usage but less efficient players (e.g., Russell Westbrook) remain above the line but fall short of similar benchmarks.

Figure 4 (NHL SDI):

- Clear positional clustering emerges, unlike in the NBA.
- Purple cluster: defensemen taking many difficult, low-percentage shots. Blue cluster: centers and wings with moderate efficiency.
- Top-left players (e.g., Brayden Point, Mark Scheifele) are mostly centers and left wings who thrive on easier shots. Right Wings (often right handed) face poorer angles and struggle near the net.
- Elite defensemen, like Cale Makar blend into offensive clusters thanks to high volume, efficient finishing, and the ability to generate offense on their own.

SDI Formulas:

SDI NBA = 0.30(Distance) + 0.20(Clock) + 0.20(Shot Type) + 0.15(Shot Zone) + 0.15(Angle)
SDI NHL = 0.40(Distance) + 0.30(Angle) + 0.20(Rebound) + 0.10(Goalie Froze)

Dashboard

Acknowledgements & Refrences

- Special thanks to our capstone instructor: Dr. Robert Buscaglia
- Datasets:
  - MoneyPuck. (n.d.). *shots_2007-2024* [Data file]. Retrieved from https://peter-tanner.com/moneypuck/downloads/shots_2007-2024.zip
  - ESPN API https://espnapi.com/