NHL Shot Statistics, Study of the Effect of Defending Ice Time on Expected Goals

Connor Raney

2024-04-30

Table of contents

1. Introduction

For my final project, I will be studying hockey data, more specifically, shot data. While there is a wide variety of publicly accessible datasets for hockey, shot data is personally most interesting to me. Modern hockey analytics rely on xGoals (expected goals), a measure of shot quality and goal probability, derived from shot data. Publicly accessible datasets contain upwards of 150 to 200 variables collected for each shot, and many of these variables are used in the calculation of xGoals. The probability of each shot becoming a goal is calculated using a model that considers multiple factors: the distance from the net, angle of the shot, type of shot, events preceding the shot, location coordinates on the ice, and numerous other variables contained in this dataset.

For context, the mean xGoals per shot measured over the last 2 seasons is approximately 0.0727, and teams average around 30 shots per game[1], which results in 2 to 3 goals per team every game. With this small value, understanding the subtle edges and effects that impact our understanding of the game becomes crucial. By identifying even minimal impacts of variables on xGoals, teams can optimize their play structure and improve their defensive game through

[1] https://www.statmuse.com/nhl/ask/nhl-team-average-shots-on-goal-per-game-2024

measurements such as the effect of defending team's ice time on xGoals. By aggregating these shot probabilities during a game, teams can calculate their expected goals to measure both offensive and defensive performance. These statistics reveal the underlying narrative of each game. They answer critical questions: Did a team underperform relative to their expected goals? Was their performance statistically justified? How well did their goalie perform? Goalie performance can be quantified by calculating the difference between expected goals and actual goals (xG - G), providing a measure of goals saved above expected. Each season, the most effective goalie typically leads in goals saved above expected (GSAAx). As evidenced, hockey statistics predominantly use the expected model as a framework, which motivated my study of factors influencing it.

My research question examines whether the collective ice time of the defensive team (defensiveTeamAverageTimeOnIce) affects the offensive team's generation of scoring chances, as measured by expected goals (xGoals). My independent variable, defensiveTeamAverageTimeOnIce, measures the average time on ice across all defensive skaters at the moment of the offensive team's shot. My dependent variable is xGoals, which I detailed earlier. While these are the primary variables in my analysis, I will incorporate various treatments and confounders throughout my analytical models.

It stands to reason that extended ice time for defensive players increases their likelihood of defensive breakdowns and conceding higher-quality chances to the offensive team. The critical question is quantifying this impact on the quality of chances conceded. This context emphasizes the importance of establishing a measurable causal relationship rather than mere correlation. The findings could significantly influence how teams deploy their players and how coaches manage ice time. The strategic implications could be substantial, depending on the demonstrated impact of this relationship on scoring chances conceded.

2. Data and Hypothesis

The data set that I have chosen to work with is Moneypuck.com's shot data dataset, which contains data with upwards of 1.7 million shots from the 2007-2008 to 2023-2024 seasons. For this analysis, I will be using only the data from the 2022-2023 and 2023-2024 seasons, as the full dataset exceeds the computational resources available for this project. The data has been downloaded from Moneypuck as of April 25, 2024, at 14:45 Eastern Time.

Here is an explanation of the dataset from Moneypuck, "There are 124 attributes for each shot, including everything from the player and goalie involved in the shot to angles, distances, what happened before the shot, and how long players had been on the ice when the shot was taken. Each shot also has model scores for its probability of being a goal (xGoals) as well as other models such as the chance there will be a rebound after the shot, the probability the shot will miss the net, and whether the goalie will freeze the puck after the shot. The data has been collected from several sources, including the NHL and ESPN. A good amount of data cleaning has also been done on the data. Arena-adjusted shot coordinates and distances are also calculated in the dataset using the strategy War-On-Ice used from the method proposed by Schuckers and Curros.[2]" As you can see, our data is observational and comes from a variety of sources, compiled into one massive dataset with many variables that can be used for analysis, which is one of the major strengths of this dataset. Moneypuck is also very straight forward with how their xGoals model works, with "factors such as the distance from the net, angle of the shot, type of shot, and what happened before the shot are key factors in the model," and you can find more information about the model on their website.[3]

The strength of this dataset lies in its comprehensive NHL coverage and pre-cleaned structure, making it well-suited to address our research question. In our analysis, the outcome variable is xGoals (labeled as xGoal in the dataset), with defensiveTeamAverageTimeOnIce serving as the treatment variable for our regression models. I will employ a propensity score matching model to measure the effect of defensiveTeamAverageTimeOnIce on xGoals, defining the treatment as values above the 75th percentile. The analysis incorporates several key covariates:

- xCordAdjusted and yCordAdjusted: absolute values of shot coordinates on the ice
- shotType: categorizing shots (slap, snap, wrist, etc.)
- shotRebound: binary indicator of rebound shots, which typically have higher goal probability
- speedFromLastEvent: measuring time elapsed since preceding events (e.g., passes or takeaways) that often lead to dangerous scoring chances
- shotAngleAdjusted: absolute value of the shot angle relative to the goal

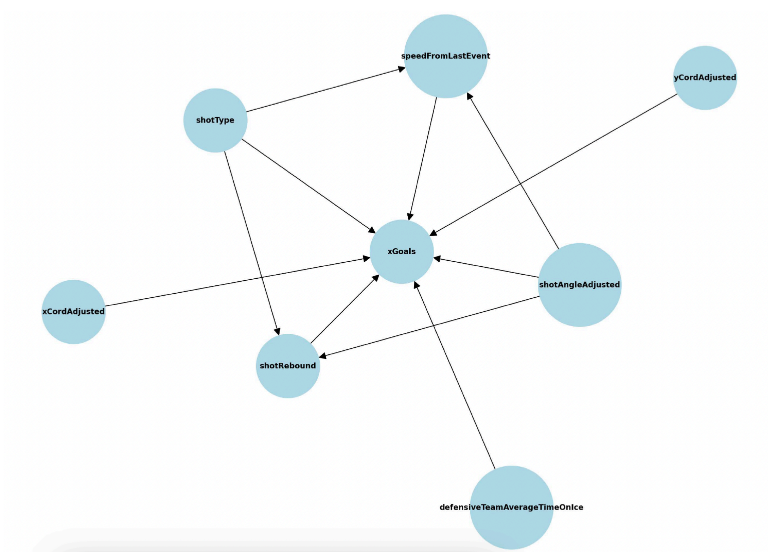

Figure 1: DAG for our X and Y variables, and our confounders & covariates

[2]Information and download for the dataset can be found at: https://moneypuck.com/data.htm
[3]Information and download for the xGoals model can be found at: https://moneypuck.com/about.htm

Drawing from extensive experience in both playing and observing hockey, I hypothesize a positive correlation between defensive ice time and xGoals conceded to the offensive team. While this relationship has received limited research attention, it aligns with fundamental hockey principles: extended ice time leads to increased player fatigue, potentially resulting in defensive breakdowns, mistakes, and reduced speed. The critical question is not whether this effect exists, but rather its magnitude. I predict a significant positive trend in xGoals as ice time increases, specifically hypothesizing a 10-20% increase in xGoals when defensiveTeamAverageTimeOnIce exceeds the 75th percentile. This hypothesis will be tested using the Propensity Score Matching model while controlling for the identified covariates and confounders.

3. Explore the Data

Initial exploration of the data revealed several noteworthy patterns. The primary analysis began with examining the distribution of xGoals relative to defendingTeamAverageTimeOnIce:

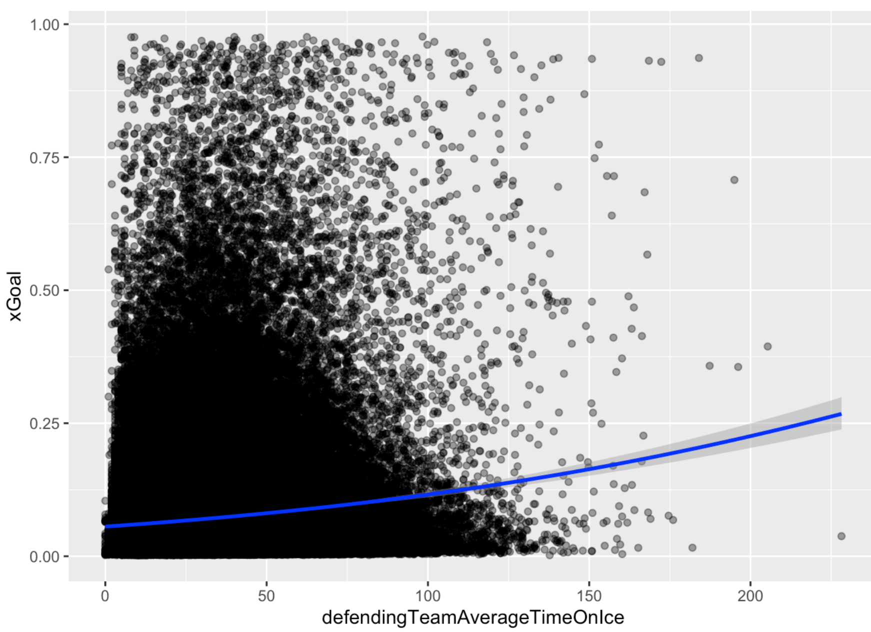

Figure 2: defensiveTeamAverageTimeOnIce over xGoals (all shots)

The distribution exhibits an unexpected pattern: while there is a general upward trend, there are numerous outliers and a cluster of small values occurring after approximately 100 seconds of ice time. Despite this clustering pattern, the data shows increasing frequency of high xGoal values as ice time increases. The logistic regression trendline displays a shallow positive slope, indicating a weak but present positive relationship in this visualization.

Examination of central tendencies revealed that the average xGoals per shot is 0.07271334, lower than the anticipated value of approximately 0.1 per shot. The distribution pattern also deviated from expectations, showing an irregular pattern of rising and then declining values with high outliers, rather than the expected steady increase over time.

An important analytical consideration emerged regarding game situations. The original dataset lacked a specific variable for game states, necessitating the creation of situation filters to compare even strength (5v5) play with non-even strength situations (5v4 power play, 5v3, or 6v5 with pulled goalie). Analysis across these three categories - even strength shots, non-even strength shots, and all shots combined - yielded significant findings. The regression analysis (shown below) indicates that non-even strength situations, particularly power plays, demonstrated a notably stronger relationship between defensive ice time and xGoals compared to even strength situations. While this situational analysis was limited in scope, it suggests potential for further investigation into the relationship between game states and xGoals calculations.

DefendingTeamAllSituations	DefendingTeamEvenStrength	DefendingTeamNonEvenStrength
-2.82896***	-2.83252***	-2.44598***
(0.01689)	(0.02136)	(0.02853)
0.00798***	0.00356***	0.00875***
(0.00041)	(0.00056)	(0.00059)

Figure 3: Linear Regressions for Each Situation (with defendingTeamAverageTimeOnIce)

Several data preparation steps were necessary before proceeding to the propensity score model. The shotType variable required conversion to a factor type, and the yCordAdjusted variable needed transformation to absolute values, similar to the pre-adjusted xCordAdjusted variable.

Descriptive statistics for the numerical variables serving as confounders and covariates in the Propensity Score Matching model revealed some unexpected patterns. Most notably, rebound shots comprised only 7% of total shots, substantially lower than the anticipated 20-30%, despite the possibility of multiple rebounds per scoring chance. Other variables showed distributions within expected ranges:

```
$xCordAdjusted
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
   0.00   49.00   65.00   61.88   78.00  100.00

$yCordAdjusted
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
   0.00    6.00   14.00   15.84   25.00   43.00

$shotRebound
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
0.00000 0.00000 0.00000 0.07156 0.00000 1.00000

$speedFromLastEvent
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  0.000   2.113   4.800   8.467  10.617 193.747

$shotAngleAdjusted
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
   0.00   16.39   30.96   33.03   46.40   88.45

$xGoal
    Min.  1st Qu.   Median     Mean  3rd Qu.     Max.
0.001537 0.014789 0.036877 0.072713 0.087609 0.976651
```

Figure 4: Descriptive Statistics for Numerical Covariates/ Confounders

4. Our Scenario - Propensity Score Matching

The Propensity Score Matching (PSM) analysis incorporates six variables to account for various effects contributing to xGoals:

1. Shot Location (x): xCordAdjusted (absolute value)
2. Shot Location (y): yCordAdjusted (absolute value)
3. Shot Type: shotType
4. Shot Rebound: shotRebound
5. Speed From Previous Event: speedFromLastEvent
6. Shot Angle: shotAngleAdjusted (absolute value)

Initially, shotDistance was considered but ultimately excluded due to its high correlation with xCordAdjusted. The first iteration of the PSM model used median defensiveTeamAverageTimeOnIce as the treatment variable; however, this resulted in a matched set nearly identical to the original dataset, yielding insignificant results. The final model employs the 75th percentile (44 seconds of defensiveTeamAverageTimeOnIce) as the treatment threshold, producing a matched dataset approximately half the size of the original, incorporating shots from all game situations. The balance and results of the treatment and control groups before and after matching are illustrated below:

	xCordAdjusted	yCordAdjusted	shotType	shotRebound	speedFromLastEvent	shotAngleAdjusted
(Intercept)	61.411***	16.264***	6.330	0.068***	8.737***	33.245***
	(0.045)	(0.027)		(0.001)	(0.026)	(0.051)
treatment	1.837***	−1.680***	−0.086	0.015***	−1.061***	−0.856***
	(0.090)	(0.054)		(0.001)	(0.052)	(0.100)
Num.Obs.	238304	238304	238304	238304	238304	238304
R2	0.002	0.004		0.001	0.002	0.000
R2 Adj.	0.002	0.004		0.001	0.002	0.000
AIC	2081528.1	1841363.0		29961.4	1822386.7	2134197.1
BIC	2081559.2	1841394.2		29992.5	1822417.8	2134228.2
Log.Lik.	−1040761.043	−920678.525		−14977.688	−911190.353	−1067095.531
RMSE	19.08	11.53		0.26	11.08	21.30
+ p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001						

Figure 5: Raw differences between treatment and control group

	xCordAdjusted	yCordAdjusted	shotType	shotRebound	speedFromLastEvent	shotAngleAdjusted
(Intercept)	63.337***	14.488***	6.236	0.084***	7.877***	32.321***
	(0.075)	(0.045)		(0.001)	(0.042)	(0.088)
treatment	−0.089	0.096	0.008	−0.001	−0.201***	0.068
	(0.106)	(0.064)		(0.002)	(0.059)	(0.125)
Num.Obs.	121062	121062	121062	121062	121062	121062
R2	0.000	0.000		0.000	0.000	0.000
R2 Adj.	0.000	0.000		0.000	0.000	0.000
AIC	1050149.1	926600.7		32099.5	908012.9	1088255.5
BIC	1050178.2	926629.8		32128.7	908042.1	1088284.6
Log.Lik.	−525071.567	−463297.346		−16046.769	−454003.473	−544124.757
RMSE	18.51	11.11		0.28	10.29	21.66
+ p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001						

Figure 6: Differences between treatment and control group, Matched Sample

The initial analysis revealed poor balance, with covariate distributions varying significantly between treatment and control groups. All variables except shotType showed significant p-values, with substantial differences in covariate coefficients. Post-matching analysis demonstrates improved balance between groups, with only speedFromLastEvent maintaining statistical significance, albeit with a considerably reduced coefficient compared to pre-matching values. The matching process effectively addressed many of the initial imbalances, resulting in better overall equilibrium between treatment and control groups. The effects estimated from both the original observational data and the PSM sample are presented below:

	Naive	Naive + Controls
(Intercept)	0.069***	-0.004*
	(0.000)	(0.002)
treatment	0.016***	0.009***
	(0.000)	(0.000)
xCordAdjusted		0.002***
		(0.000)
yCordAdjusted		-0.001***
		(0.000)
shotTypeBACK		-0.022***
		(0.002)
shotTypeDEFL		-0.012***
		(0.002)
shotTypeSLAP		-0.012***
		(0.002)
shotTypeSNAP		-0.009***
		(0.002)
shotTypeTIP		-0.018***
		(0.002)
shotTypeWRAP		-0.037***
		(0.003)
shotTypeWRIST		-0.014***
		(0.002)
shotRebound		0.110***
		(0.001)
speedFromLastEvent		0.001***
		(0.000)
shotAngleAdjusted		-0.001***
		(0.000)
Num.Obs.	238304	238304
R2	0.005	0.338
R2 Adj.	0.005	0.338
AIC	-410737.7	-507925.5
BIC	-410706.5	-507769.8
Log.Lik.	205371.846	253977.772
RMSE	0.10	0.08

+ p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

Naive Models

	Effects	Effects + Controls
(Intercept)	0.077***	-0.015***
	(0.000)	(0.003)
treatment	0.008***	0.008***
	(0.001)	(0.001)
xCordAdjusted		0.002***
		(0.000)
yCordAdjusted		-0.001***
		(0.000)
shotTypeBACK		-0.023***
		(0.003)
shotTypeDEFL		-0.017***
		(0.003)
shotTypeSLAP		-0.012***
		(0.003)
shotTypeSNAP		-0.009**
		(0.003)
shotTypeTIP		-0.023***
		(0.003)
shotTypeWRAP		-0.039***
		(0.004)
shotTypeWRIST		-0.013***
		(0.003)
shotRebound		0.107***
		(0.001)
speedFromLastEvent		0.001***
		(0.000)
shotAngleAdjusted		-0.001***
		(0.000)
Num.Obs.	121062	121062
R2	0.001	0.301
R2 Adj.	0.001	0.301
AIC	-186609.8	-229765.2
BIC	-186580.7	-229619.6
Log.Lik.	93307.891	114897.576
RMSE	0.11	0.09

+ p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

Effects w/PSM

Figure 7: Estimate with and without Propensity Score Matching

The PSM results reveal a substantial treatment effect, with an 11% increase in xGoals for the treatment group ($p < .001$). The treatment coefficient's progression—decreasing from 0.016 to 0.009 in the naive model with controls, then stabilizing at 0.008 in both PSM models—indicates robust treatment effects across different modeling approaches. Post-PSM analysis shows continued significance of covariates while maintaining treatment consistency, suggesting effective control of confounding variables through the matching process. The model's explanatory power improves significantly with the inclusion of covariates, as evidenced by R^2 increases from 0.005 to 0.338 in the naive model with controls, and from 0.001 to 0.301 in the PSM model with controls. These improvements indicate that the covariate-adjusted models explain a substantially larger proportion of variance in xGoals. The PSM model successfully demonstrates the consistent positive relationship between defensiveTeamAverageTimeOnIce and xGoals, with results remaining stable across both naive and PSM approaches when controlling for covariate effects.

5. Conclusions

This study addressed the question: Does the amount of time the defensive team (defensiveTeamAverageTimeOnIce) has collectively been on the ice affect the other (offensive) team's generation of scoring chances, measured by expected goals (xGoals)? Analysis through both naive models and PSM models, supported by effective matching processes, demonstrates that defensiveTeamAverageTimeOnIce significantly affects the offensive team's generation of scoring chances. The results show an 11% increase (0.008 xGoals) in expected goals for shots above the 75th percentile of defensiveTeamAverageTimeOnIce (44 seconds)—a substantial increase given the average xGoals per shot of 0.0727. In hockey, where marginal advantages accumulate over time, these findings have important implications for game management. The data suggests that keeping players on ice beyond 44 seconds (compared to the average of ~32 seconds) represents a measurable defensive liability. This evidence supports conventional wisdom favoring shorter shifts (around 30 seconds) and suggests that teams might benefit from even stricter ice-time management to maintain defensive effectiveness.

The research process yielded insights beyond the primary findings. The implementation of propensity score matching proved particularly instructive, highlighting the importance of proper data subsetting in the matching process and its impact on analytical outcomes. A significant limitation emerged from computational constraints—while the complete dataset contained 1.7 million shots, processing limitations of an M1 Mac restricted analysis to approximately 200,000 shots from two seasons, rather than the intended 10+ seasons. Access to more substantial computing resources, such as cluster or cloud computing, would likely strengthen the findings' confidence levels. This limitation raises questions of external validity: whether the observed patterns would persist across the full 1.5 million shots remains uncertain. Additionally, internal validity considerations arise from the numerous confounding variables affecting xGoals measurement, not all of which could be accounted for in the current model.

This research underscores an important methodological lesson: while perfect answers may remain elusive, carefully constructed models can provide valuable insights within acknowledged limitations. Future research opportunities include expanding the analysis to the complete dataset, refining the models, and potentially incorporating additional variables. The project successfully merged statistical methodology with practical hockey analytics, producing findings with potential applications for game strategy and player management.