Analyzing Scoring Chance Events Created by Passes in Women's Ice Hockey

Andy McBurney, Dartmouth College
Quantitative Social Science One-Quarter Project

Introduction

- Prior to the Big Data Cup 2021, passing data in ice hockey was not publicly available.
- With this new data, new insights can be made about the importance of passing in ice hockey, how different passes create scoring chances, and how ice hockey is played in general.
- Previous research has looked into creating different expected goals, expected assists, expected threat, etc metrics, but has not taken into account different passing types.
- Most of the highest-quality scoring chances happen in ice hockey after a pass is made across the midline of the ice ("royal road" pass), after a pass is made from behind the goal line to in front of the net, and off of one-timer shots.
- One-timers can be very effective as a one-timer can be shot even harder than a slapshot by the most skilled players, and since the shot is taken immediately after a pass, the shooter can often shoot before the goalie has moved into the correct position to stop the shot.
- Despite the advantage of one-timers, they a very high-skill shots, and often players will fan, not hit the puck correctly, have their shot blocked, or miss the net due to the difficulty of correctly shooting a one-timer
- Previous research in the NHL has found that one-timers taken in "the house", or the area where the highest-quality shots are taken close to the net near the midline of the ice, are very effective
- Other research has found that one-timers taken not in "the house" have extremely low scoring rates, and thus an inverse effect of the one-timer when it is taken not in "the house" compared to non-one-timers
- There is no current research into the effectiveness of one-timers in women's ice hockey
- Considering women's ice hockey tends to be a slower game than men's ice hockey and has slightly different rules, I will attempt to determine if the effects of taking a one-timer in women's ice hockey are similar to the effects of taking a one-timer in men's ice hockey.

Questions and Hypotheses

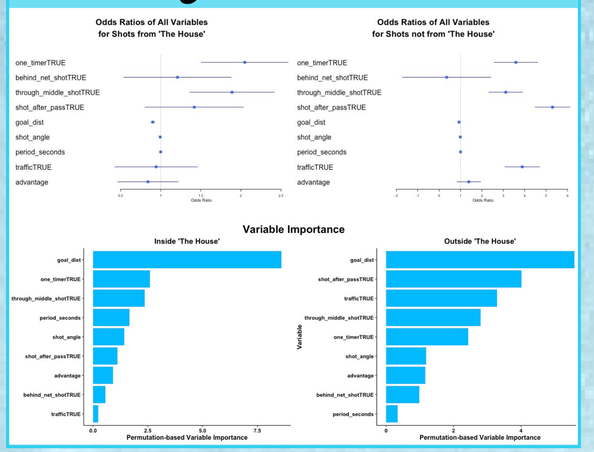
1. How different are the shooting percentages of one-timers in women's hockey, as compared to non-one-timers?
   a. H1: One-timer shooting percentages in women's hockey will be comparable to men's hockey
2. How big of a difference does a pass through the midline of the ice before a shot or a pass that originates behind the net affect the shooting percentages of one-timers and non-one-timers?
   a. H2: Both passes will create better scoring opportunities and higher shooting percentages, though passes through the middle of the ice will create better scoring opportunities than passes from behind the net.
3. Is there a similar relationship in women's ice hockey as in men's ice hockey between one-timer shots and shooting percentage from inside and outside "the house"?
   a. H3: Yes, and one-timer shots for women outside "the house" will be higher shooting percentages than men's hockey.

Data and Methods

- Data from 17 Olympic, 2 NCAA, and 15 PHF games in 2018-2022 with 13 total teams.
- 168 goals, of which 39 goals were one-timers, 34 goals were directly after receiving a pass, and 95 goals were neither one-timers nor quick goals after a pass
- 4,360 shot attempts, of which 551 shot attempts were one-timers, 636 shot attempts were directly after receiving a pass, and 3,173 were neither one-timers nor quick shots after a pass
- Model: Multivariable logistic regression

$$xGoal = \beta_0 + \beta_1\, one\,timer + \beta_2\, behind\,net\,shot + \beta_3\, through\,middle\,shot + \beta_4\, shot\,after\,pass + \beta_5 goal\,dist + \beta_6 shot\,angle + \beta_7 shot\,angle + \beta_8 traffic + \beta_9 advantage + \beta_{10} period\,seconds + \epsilon$$

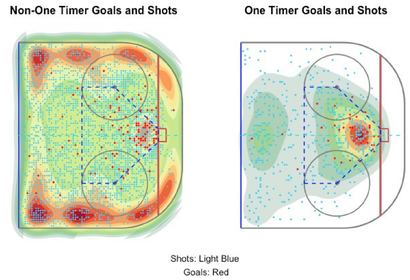
Regression Results



Figures and Tables

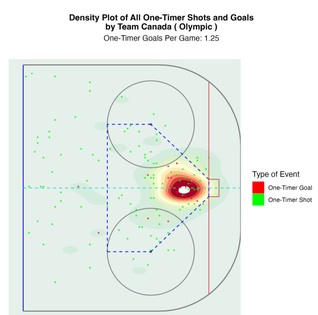
All Teams Density Plot

2D Density plots of Non-One Timer Shots & Goals vs. One-Timer Shots & Goals

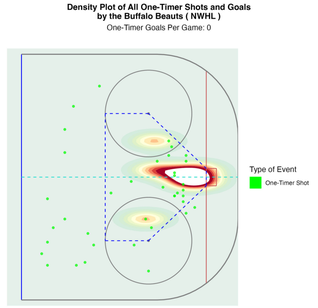


Shots: Light Blue
Goals: Red

Best One-Timer Team Density Plot



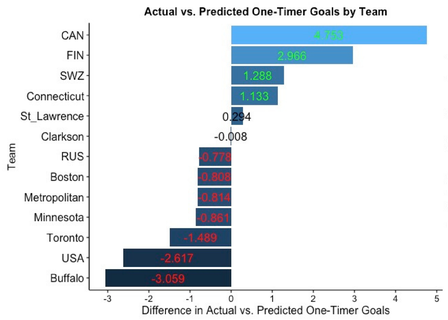
Worst One-Timer Team Density Plot

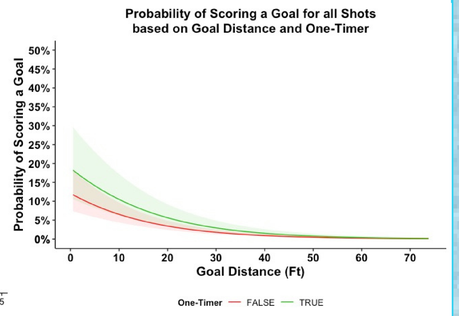


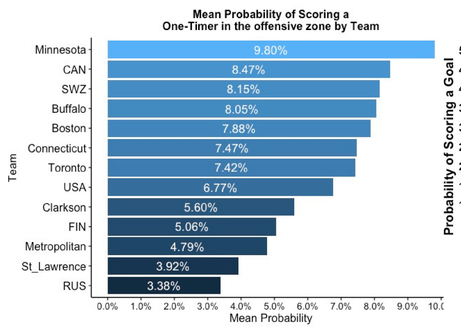
Difference in Shotting Percentage of events from "The House" and not from "The House"

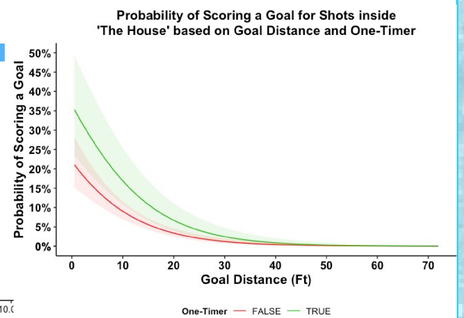
One-Timer	Shot After Pass	Shot after pass through midline	Shot after pass from behind the net	Shooting Percentage Difference
FALSE	FALSE	FALSE	FALSE	6.07
FALSE	FALSE	FALSE	TRUE	1.85
FALSE	FALSE	TRUE	FALSE	2.41
FALSE	FALSE	TRUE	TRUE	0.00
FALSE	TRUE	FALSE	FALSE	2.87
FALSE	TRUE	FALSE	TRUE	0.00
FALSE	TRUE	TRUE	FALSE	8.17
FALSE	TRUE	TRUE	TRUE	28.57
TRUE	FALSE	FALSE	FALSE	4.98
TRUE	FALSE	FALSE	TRUE	15.62
TRUE	FALSE	TRUE	FALSE	16.04
TRUE	FALSE	TRUE	TRUE	11.11

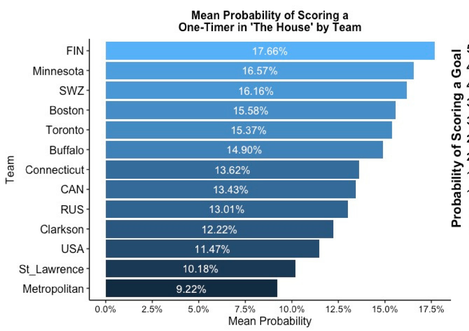
One-Timer Scoring by Team

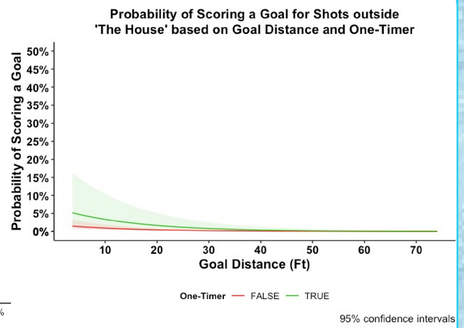




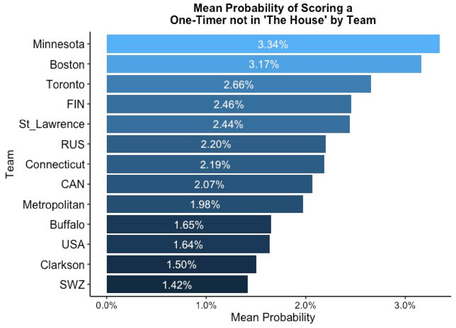


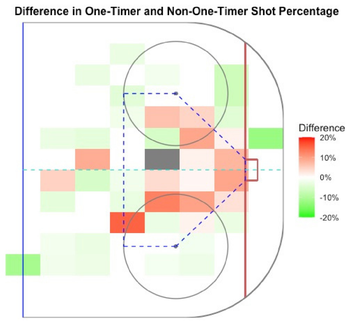






95% confidence intervals





Results

Shots from Inside "The House"

- One-timers significantly increase the chances of scoring a goal.
- One-timers were the second most important variable in the logistic regression, behind goal distance.
- Shots that were taken after a pass through the middle of the ice ("royal road" passes) also significantly increase the chances of scoring a goal.
- The further a shot is taken from the net, the lower the chances of scoring a goal.
- Whether a shot was taken after a pass from behind the goal line, whether a shot was taken quickly after receiving a pass, the angle of the shot to the net, whether there was traffic in between the shooter and the goalie, and whether the shooter was on a power play or not, and the amount of time left in the period did not significantly change the chances of scoring a goal.

Shots from Outside "The House"

- One-timers significantly increase the chances of scoring a goal, though one-timers outside "the house" have a smaller effect on whether a shot will result in a goal than one-timers inside "the house".
- One-timers were not nearly important in the regression for outside "the house" compared to inside "the house"
- Shots that were taken after a pass through the middle of the ice ("royal road" passes) also significantly increase the chances of scoring a goal.
- The further a shot is taken from the net, the lower the chances of scoring a goal.
- Whether a shot was taken after a pass from behind the goal line, whether the shooter was on a power play or not, and the amount of time left in the period did not significantly change the chances of scoring a goal.
- Shots directly after a pass and traffic in between the shooter and goalie had a significant effect on predicted probability of scoring.

Team Performances

- Team Canada consistently scored more one-timers than predicted.
- Minnesota creates a lot of one-timer chances and converts those chances near their predicted shooting percentage for one-timers.
- Despite taking quite a few one-timers in "the house" and near the goal, Buffalo was not able to score on any of their chances.
- Even the best team and players predicted probability of scoring a one-timer outside "the house" is low, and not significantly higher than the predicted probability of scoring a non-one-timer.
- Comparing all the possible shot types and passes before a shot, the largest differences in shooting percentages between plays inside and outside "the house" are from one-timers, where one-timers in "the house" have considerably higher conversion rates while one-timer conversion rates outside "the house" are low.

Discussion & Conclusion

Key Findings

- Goal distance is the most important factor when predicting goals
- Shots further than 30 ft away from the net have very low chances of scoring.
- Knowing whether a shot is a one-timer or not improves shooting percentage predictions, especially for shots in "the house".
- The multivariable logistic regression model was 72% accurate for shots in "the house", 83% accurate for shots outside "the house", and 73% accurate for shots overall

Study Decisions

- Shots that were blocked or missed were kept in the dataset, which may reduce the shooting percentage of shots further away as they have a higher chance of missing the net or being blocked.
- Passes from the neutral zone into the offensive zone were not included.
- A logistic regression model was used; other models such as random forest or extreme gradient boosting models have been shown to be more accurate than linear or logistic regression models.
- Only the final pass before a shot were included; other research has shown that including multiple plays leading up to a shot in a model improves its predictive accuracy.
- The type of shot for non-one-timers was not considered in this model
- Pass distance was not considered in this model.

Limitations

- The sample size in terms of the number of games, shots, goals, and teams was limited, making it difficult to find statistical significance.
- Other information on the pass were missing, such as the speed of the pass, whether the pass was along the ice or in the air, if the pass was easy to receive, etc.
- Other information on shots were missing, such as shot speed, shot location in the net, shooter ability, etc.
- There was not enough data to analyze predicted probabilities of goals on a player level.