A Puck Above the Rest: Exploring the Effects of New Data on 2020 NHL Draft Decisions

Ashley Mullan and Lucy Ward

October 25, 2020

Advisors: S. Ventura, N. Citrone, R. Yurko
The NHL draft usually runs in late June.

COVID-19 forced the 2020 draft to run in early October.

Multiple European leagues began prior to the draft, so 2020 prospects from these leagues had more data available.

Objective: Model players’ future performance given their amateur performance and assess the impact the of additional data on their value
Data

- Season-level data from amateur and professional seasons from 2010 to 2020
- Player characteristics and statistics
- Team-level statistics (games played, goals scored, goals against)
- New metrics (relative age, PTPP)

<table>
<thead>
<tr>
<th>ID</th>
<th>Name</th>
<th>DOB</th>
<th>Country</th>
<th>HT</th>
<th>WT</th>
<th>Position</th>
<th>Shoots</th>
<th>GP</th>
<th>G</th>
<th>A</th>
<th>P</th>
<th>PM</th>
</tr>
</thead>
<tbody>
<tr>
<td>199870</td>
<td>Jake Guentzel</td>
<td>1994-10-06</td>
<td>USA</td>
<td>180</td>
<td>82</td>
<td>LW/C</td>
<td>L</td>
<td>60</td>
<td>29</td>
<td>44</td>
<td>73</td>
<td>13</td>
</tr>
</tbody>
</table>
Data Modification and Response Metric

- Player statistics scaled by number of games played
- Response Metric: Professional Total Point Percentage (PTPP)

\[
\text{Player Points per Game} = \frac{\text{player's total season points}}{\text{player's total games in season}}
\]

\[
\text{Pro Team Points per Game} = \frac{\text{pro team's total season points}}{\text{games in season}}
\]

\[
\text{PTPP} = \frac{\text{Player Points per Game}}{\text{Pro Team Points per Game}}
\]
Prior Research

- Earlier draft choice results in better outcomes, but this effect is muted after 100 decisions. (Tingling et al, 2011)

- A Poisson GAM successfully models time on ice while including non-linear effects. (Schuckers, 2016)

- Players born in the first quarter made up the greatest percentage of the draft class. (Deaner et al, 2013)

- Early birthdays, size advantages, and anaerobic power increase chances of draft selection. (Rocznioe et al, 2013)
Approach

- Identify players from target leagues, both those who were drafted to the NHL and those who were not
- Model future performance given selection to NHL
- Scale expected performance by probability of making NHL
- Observe how predictions change when more data is added to simulate early Fall 2020 games
Target Leagues: Early Season Starts

- Czech Leagues
  - (Czech, Czech2)
- Russian Leagues
  - (KHL, MHL)
- Swiss Leagues
  - (NLA)
- Swedish Leagues
  - (SHL, Allsvenskan, J18-Allsvenskan, Superelit)
Where are new players getting drafted from?

2010-2018 NHL Drafts

Number of Players SuccessfullyDrafted
Eliminating Players With Fewer Than 20 Career Games
Draft Probability Model

Goal: Predict a player’s probability of being drafted into the NHL.
Most Players Have a Very Small Probability of Being Drafted

2018 Draft Class Prospects from Target Leagues

![Chart showing most players have a very small probability of being drafted.](chart.png)
NHL Performance Model

Goal: Given that a player was drafted into the NHL, model his value to his NHL team based on his amateur statistics.
Challenges in Predicting Professional Performance

Comparison of Holdout Season Error

- 2010
- 2011
- 2012
- 2013
- 2014
- 2015
- 2016
- 2017

Holdout Season Predictions vs. Holdout Season Observed
Joint Metric: Combining Draft Probability with Expected Performance

Two Seasons are Better than One
Weighted Joint Metric Value Incorporating New Season Data (2019 Class)
Future Directions

- Improve accuracy of performance model with a potential different response variable
- Consider how to correct for lack of independence of draft probability and expected performance
- Incorporate league strength metric into player assessment to allow for broader application
Any questions?

ashley.mullan@scranton.edu
@ashley___mullan

lward7@uwyo.edu
@_lucyward_