UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

PILOT EVALUATION OF THE WEIGHTED OFFENSIVE PRODUCTIVITY RATING IN HOCKEY USING LONGITUDINAL DATA AND LOGISTIC REGRESSION

A THESIS

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

Degree of

MASTER OF SCIENCE

By

JARROD BLUMER
Norman, Oklahoma
2018

PILOT EVALUATION OF THE WEIGHTED OFFENSIVE PRODUCTIVITY RATING IN HOCKEY USING LONGITUDINAL DATA AND LOGISTIC REGRESSION

A THESIS APPROVED FOR THE
DEPARTMENT OF HEALTH AND EXERCISE SCIENCE

BY

Dr. Daniel Larson, Chair

Dr. Rebecca Larson

Dr. Christopher Black

Acknowledgements

I would like to express my sincere gratitude to my advisor and committee chair Dr. Daniel Larson, as well as my fellow committee members Dr. Rebecca Larson and Dr. Christopher Black for providing guidance and recommendations throughout the entire thesis process. I'd like to thank the research team of undergraduate research assistants for grinding through countless hours of film during data collection.

I would also like to thank Dr. John Clark for telling me this was an idea worth researching back when it was only scribbles on a notepad during undergraduate school.

Lastly, I would like to my family, as without them I would not be here today.

Table of Contents

List of Tables

List of Figures

Abstract

The purpose of this study was to evaluate an analog to the baseball statistic Weighted On-Base Average (wOBA), which provides an estimate of a baseball player's offensive production, in the context of hockey. In this model, a hockey player's offensive skillset is determined by passing, zone-transition, and shooting, which together create a Weighted Offensive Productivity Rating (wOPR). The contribution of these components were estimated using shot assist, shot attempt, zone entry, and zone exit data, with linear weights produced separately for forwards and defensemen. Logistic regression estimations were used to analyze whether the wOPR is a better predictor of game success compared to a more traditional measure such as a player's total points in NCAA Division I Hockey. Each prediction equation consisted of the performance metric, the quality of competition, and a quality of teammates' metric for any given player in a given game. The wOPR at the team level in 5 on 5 play is a statistically significant predictor of game results. The wOPR appears to perform better as an indicator of play quality in instances where data availability about individual players may be limited, e.g. new players or early season play. Preliminary analysis supports the adoption of the model in practice.

Chapter 1

Introduction

Data analytics is the analysis of large sets of data through the use of mathematics, statistics, and computer software to identify patterns and other meaningful information (Dictionary.com, 2018). Baseball has done this for quite some time and the implementation of analytics has recently gained traction in hockey (Mondello & Kamke, 2014). However, there still are limited advanced metrics to aid in the player evaluation process (Macdonald, 2011). Macdonald (2011) mentions Tom Awad's Goals Versus Threshold, Alan Ryder's Player Contribution, and also Timo Seppa's Even-Strength Total Rating as a few of the previous attempts to measure player contributions, in addition to his own adjusted plus-minus metric. Given this issue, it is important to further develop advanced statistical models or metrics to give hockey organizations more tools to evaluate player performance. Perhaps the biggest obstacle hindering acceptance of these progressive ideas is the data currently available to decision makers and the public. While one could assume that few NHL teams have access to proprietary information to make decisions, many are left to use basic box-score statistics and large-scale change may not happen until the implementation of player and puck tracking technologies (Wheeldon, 2017).

Used as the foundation for the offensive component of Wins Above Replacement (WAR), Weighted On-Base Average (wOBA) attempts to measure a hitter's offensive production based on relative values of specified offensive outcomes (Fangraphs). According to Baumer et al. (2015), wOBA gives one of the better representations of a player's offensive production as it values the possible offensive outcomes for a hitter's plate appearances proportionally to the likelihood the outcome equates into a run scored in baseball.

In an effort to adapt the Weighted On-Base Average metric to valuing a hockey player's offensive productivity, one can consider what makes up the offensive portion of a hockey player's skillset when he has possession of the puck. Based on previous studies and background research, we suggest that a hockey player's offensive production during even-strength play[1] is a result of his passing, shooting, and zone transition abilities, which in this model together create a Weighted Offensive Productivity Rating (wOPR). These 3 components are composed of specific on-ice events that attempt to measure each component accurately and in their entirety. These events include secondary shot assists, primary shot assists, shot attempts, zone entries, and zone exits. In addition to representing their corresponding component, the measured variables can also be used to create linear weights. In order to scale the relative importance of each type of event, each of the passing, transition, and shooting components' weights are calculated using the proportion of time each event historically leads to a goal. Additionally, these linear weights are calculated separately for Forwards and Defensemen, as it would be wrong to assume forwards and defensemen to produce offensively at identical rates (Macdonald, 2011).

The contribution of these components are estimated using shot assist, shot attempt, zone entry, and zone exit data from Robert Morris University Men's Division I Ice Hockey team's 2016-2017 season. The Weighted Offensive Productivity Rating

[1] Even-strength play in this study consists of game play when each team has 5 skaters on the ice, excluding all other strength states such as power-play and penalty-kill situations.

model is used to determine which players are most productive, offensively, when the puck is in their possession. Using the previously mentioned outcomes and appropriate linear weights, we may be able to establish a strong foundation for a useful Wins Above Replacement statistic or similar catch-all statistic for player evaluation in hockey at all levels.

Finding accurate player evaluation metrics for hockey players has been an issue in the past as teams often used simple counting statistics, such as points and plus-minus, for evaluation and contractual decision-making. The purpose of this study is to quantify the offensive productivity of hockey players based on their passing, shooting, and zone transition abilities. The ideal end result will be a statistical model with increased predictive power and limited variability over the course of the season.

Purpose of Study

The purpose of this study is to evaluate the Weighted Offensive Productivity Rating (wOPR) relative to a traditional evaluation metrics such as player total points, specifically in a limited data availability situation.

Research Questions

1. Is the Weighted Offensive Productivity Rating a valid indicator of on-ice performance of hockey players at the individual player level?
2. Is the Weighted Offensive Productivity Rating a valid indicator of on-ice performance of hockey players at the team level?
3. Is the Weighted Offensive Productivity Rating a more accurate representation of performance than traditional Points metrics in small dataset situations?

Hypotheses

Research Hypotheses

1. The Weighted Offensive Productivity Rating is a valid indicator of on-ice performance of hockey players at the individual player level.
2. The Weighted Offensive Productivity Rating is a valid indicator of on-ice performance of hockey players at the team level.
3. The Weighted Offensive Productivity Rating more accurately represents player performance than traditional Points metrics in small dataset situations.

Null Hypotheses

1. The Weighted Offensive Productivity Rating model shows no relationship to on-ice performance of hockey players at the individual player level.
2. The Weighted Offensive Productivity Rating model shows no relationship to on-ice performance of hockey players at the team level.
3. The Weighted Offensive Productivity Rating shows no relationship to player performance than traditional Points metrics in small dataset situations.

Significance of Study

For much of the sport's existence, hockey organizations and their decision makers have typically used simple counting statistics, such as goals, assists, plus-minus (a player's goal differential when he is on the ice during even strength play), and qualitative assessment to evaluate the on-ice performance of their players. While baseball was an early adopter of robust quantitative performance assessment, only recently have hockey organizations and decision makers begun to use newer and more statistically advanced methods to evaluate player performance (Mondello & Kamke, 2014). The proposed Weighted Offensive Productivity Rating model would add to and improve the few publicly available metrics that attempt to distill multiple areas of a hockey player's skillset in a single measure. This is quite similar to the way the Weighted On-Base Average statistic is used for evaluating baseball players hitting abilities. This exact method has yet to be attempted in hockey, and while there have been other proposed player evaluation models for hockey, most are yet to be widely accepted in the way that WAR and wOBA have been accepted in baseball. A frequent issue that has arisen with these previously proposed models is, not long after preliminary analysis is made available to the public, metric developers are plucked from the public sphere by sport organizations and future modifications and analyses are made private. This results in the public being forced to decide among themselves what metric(s) are most accurate.

Additionally, data at the NCAA level is suboptimal. While data at the NHL level has been made publicly available via resources such as war-on-ice.com in the past and most recently corsica.hockey, these would not be accessible without the work of individuals in the public hockey analytics community web-scraping through thousands of NHL play-by-play files (Thomas, 2014; Perry, 2017). Moreover, even the NHL data that is available to the public has limitations in its' own right. This type of work is incapable of being done at the NCAA level due to the underwhelming play-by-play files. The implementation of Weighted Offensive Productivity Rating model will potentially solve this shortcoming, and give the hockey community a Weighted On-Base Average-like metric to evaluate players with greater accuracy and predictive power.

Hypothetically speaking, an NCAA coach could be in a situation where 4 games into the season an injury to a star player leaves a glaring hole in the team's lineup, and said coach is now forced to determine who will fill that position. The coach would ideally replace the injured player with a player that is capable of providing the team with as much overall production as possible. However, being only 4 games into the season, the players the coach is considering have accumulated little to no points (goals and assists). The wOPR would allow the coach to objectively select the player who offers the most offensive productivity to this point in the season. As opposed to making the decision based on cognitive bias, subjective ideas, and insignificant counting statistics, the wOPR provides the coach with a tool to differentiate between the players in this small dataset situation.

Delimitations

The delimitations of this study include:

1. This study uses the NCAA Division I hockey players from Robert Morris University in Pittsburgh, Pennsylvania.
2. This study includes 23 skaters in total with 8 defensemen and 15 forwards.
3. Data has been collected from 34 games during the 2016-2017 season.
4. Data has been collected during 5 on 5 strength state.

Limitations

The limitations of this study include:

1. Data was being collected from 1 team from 1 league.
2. This is a convenience sample due to access to their game film only.

3. Data was being manually collected by 4 data trackers and may be open to scorer variability.
4. Not all skaters played in all 34 games.

Assumptions

The assumptions of this study include:

1. Players are not attempting to artificially inflate the variables that are being used.
2. The video coordinator properly recorded the events and did not systematically omit or miss-code a significant number of events.
3. All data trackers correctly recorded each sequence of events in their entirety.
4. The recorded events appropriately represent a player's offensive abilities.

Operational Definitions

For several of these operational definitions, the image in Figure 1 was used from the War-On-Ice hockey analytics website. The graphic provides a framework for determining the difference between low, medium, and high danger shot attempts, where attempts close to the net (light grey) have a higher probability of resulting in a goal and shot attempts from distance (lightest grey) have the lowest probability of resulting in a goal.

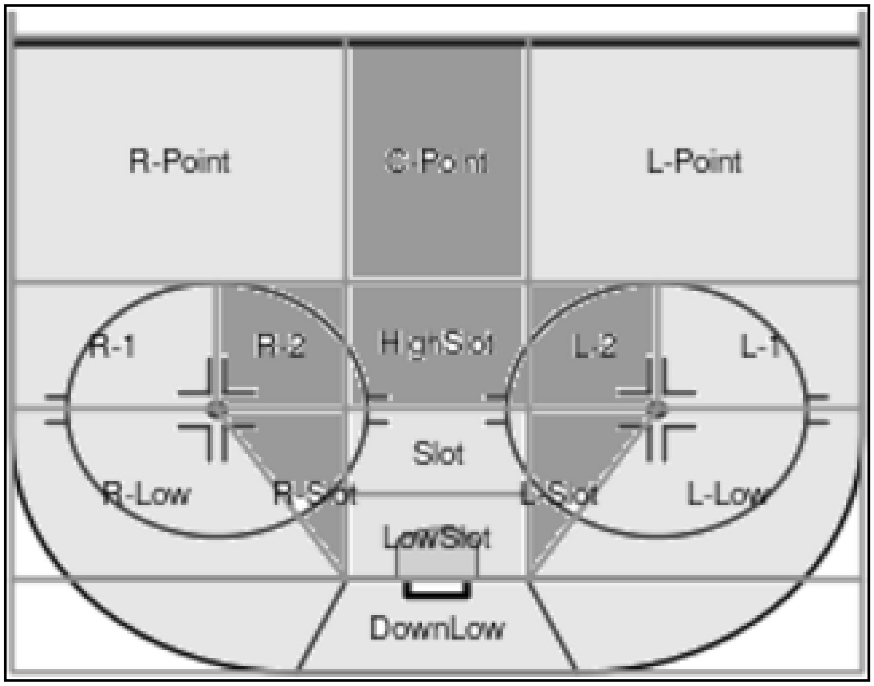

Figure 1. Shot Danger Map. Reprinted from War-On-Ice, by Thomas. A, 2014. Retrieved from war-on-ice.com.

The operational definitions of this study include:

Controlled zone entry (CE): the process of a player moving the puck from either his own defensive zone or the neutral zone and into the offensive zone by controlling the puck himself or completing a pass to a teammate

Controlled zone exit (CX): the process of a player moving the puck out of his own defensive zone by controlling the puck himself or completing a pass to a teammate

Goal (G): credited to the last player to touch the puck prior to it entering the opposing team’s net.

High-Danger Primary shot assist (HSA1): the pass that occurs prior to a high-danger shot attempt

High-Danger shot attempt (HSA): a shot on goal, blocked shot, or missed shot from a high-danger area according to danger-zone graphic from War-On-Ice

Linear weights (w): the probability that a certain event results in a goal, is calculated separately for forwards and defensemen and for each type of event

Low-Danger Primary shot assist (LSA1): the pass that occurs prior to a low-danger shot attempt

Low-Danger shot attempt (LSA): a shot on goal, blocked shot, or missed shot from a low-danger area according to danger-zone graphic from War-On-Ice

Medium-Danger Primary shot assist (MSA1): the pass that occurs prior to a medium-danger shot attempt

Medium-Danger shot attempt (MSA): a shot on goal, blocked shot, or missed shot from a medium-danger area according to danger-zone graphic from War-On-Ice

Passing: model component consisting of secondary shot assists, and all primary shot assists

Secondary shot assist (SA2): the pass that occurs 2 passes prior to a shot attempt

Shooting: model component consisting of all shot attempts

Zone Transition: model component consisting of controlled zone exits and exit assists, and controlled zone entries and entry assists

Weighted Offensive Productivity Rating (wOPR): the overall model being tested using on-ice offensive events and weighting those events based on the probability those events result in a goal

Weighted On-Base Average (wOBA): a baseball player’s offensive production as it values the possible offensive outcomes for a hitter’s plate appearances proportionally to the likelihood the outcome equates into a run scored (Baumer, Jensen, Matthews, 2015)

Wins Above Replacement (WAR): single value metric used to summarize a baseball player's contribution to his team (Baumer, Jensen, Matthews, 2015)

Chapter 2

Literature Review

The proposed and evaluated Weighted Offensive Productivity Rating model incorporates a hockey player's passing, zone transition, and shooting abilities. The primary research questions for this study were:

1. Is the Weighted Offensive Productivity Rating a valid indicator of on-ice performance of hockey players at the individual player level?
2. Is the Weighted Offensive Productivity Rating a valid indicator of on-ice performance of hockey players at the team level?
3. Is the Weighted Offensive Productivity Rating a more accurate representation of performance than traditional Points metrics in small dataset situations?

This chapter will focus on previously completed research that was used as the ideological foundation for this model. The first section will give an overview of previous examples of analytical decision-making utilized in different levels of competition in professional sports. The second section will offer further explanation of Weighted On-Base Average, while touching on key concepts of single-value metrics. The third section includes the theory used to replicate the previously mentioned baseball evaluation models in a manner that would allow them to be applicable to hockey players. The last section will examine previously demonstrated research methods and techniques used by practitioners of sport analytics, and how these approaches may be relevant, in whole or in part, when finalizing the Weighted Offensive Productivity Rating model.

The literature review process consisted of searching through several scholarly archives including EBSCOHOST, De Gruyter, and Google Scholar in an effort to find previously published research pertaining to analytics applied in professional sports, baseball's adoption of advanced quantitative evaluation methods, and previous attempts to apply statistical analysis to hockey players. Some the most frequently used keywords consisted of "hockey," "WAR," "Weighted On-Base Average," and "sport analytics," as displayed in Table 1. In addition to reviewing past studies, searches were also completed to review statistical methods and theories in an effort to adequately clarify definitions and concepts such as "descriptive correlational studies" and "logistic regression interpretations," which are displayed in Table 2. While the *Journal of Quantitative Analysis* provided ample relevant literature for this study, Google Scholar was used in an attempt to identify prior scholarly research pertaining to "hockey performance evaluation metrics" and "hockey performance metrics," however relevant research was not found.

Analytics in Professional Sports

The primary goal of a professional sport organization's front office is to effectively and efficiently evaluate personnel in order to achieve on-field success. It is important to develop a process that the decision-makers deem to be effective and efficient. Effective in this case would be defined as identifying and acquiring productive players that contribute to the team's overall success, while efficient would be the allocation of resources in a manner that extracts the most value to the organization. For sport organizations, those resources include the organization's payroll, staffing, and the time and effort decision-makers and staff spend on the evaluation process.

In order to gain a competitive advantage in the process of player evaluation and resource allocation, professional sport organizations have recently turned to the use of data analytics to aid in the process of objective decision-making (Mondello & Kamke, 2014). The most prominent example of this was detailed in *Moneyball* (Lewis, 2004), when the Oakland Athletics of Major League Baseball instituted a wide-spread adoption of analytical decision-making when it came to their player personnel. Oakland was not capable of operating under an upper-echelon payroll, and where forced to field a competitive team under strict financial constraints. The organization focused on using advanced statistical analysis when evaluating players in order to uncover strategical competitive advantages and hidden value in players that other teams did not know existed. One of the most noteworthy applications of statistical analysis by the Oakland front office was the denouncing of a player's batting average (AVG) in favor of on-base percentage (OBP%) and slugging percentage (SLG%), frequently combined into on-base plus slugging percentage (OPS), when evaluating hitters. This was one of the first attempts at using a single-value metric to value baseball players. The logic behind this theory was that a player's batting average assumed all hits were of equal value, in addition to ignoring a batter's ability to draw walks. On-base percentage addressed the issue of accounting for a player's plate discipline by including walks, while slugging percentage was baseball's first attempt to value different hits appropriately, primarily distinguishing singles from extra-base hits. The resulting combination of OPS was even further refined by a statistic known as Weighted On-base Average (wOBA).

While the adoption of OPS was just a minute part of Oakland's adoption of using data analytics for player evaluation, discovering this, in addition to all of the other competitive advantages the front office uncovered, would have been all for not if the analyses were not applied by decision-makers. How an organization utilizes data analytics is equally important as the statistical findings themselves, as an objective decision-making process can help exploit these market inefficiencies.

With a relatively widespread adoption of analytical decision making across Major League Baseball in recent years, these evaluation ideologies and practices are making their way into lower levels of competition, such as minor league baseball and collegiate baseball. An exhibit of this is expressed in *The Only Rule Is It Has to Work: Our Wild Experiment Building a New Kind of Baseball Team* (Lindbergh & Miller, 2016), an in-depth depiction of two analytically-slanted baseball writers that took over an Independent League baseball team with the goal of operating the team using strictly advanced statistical analysis. Over the course of one season Ben Lindbergh and Sam Miller employed some of the most untraditional and unconventional baseball strategies because they had data and advanced statistical analyses that supported their decisions and mathematically increased their chances of successful outcomes. The biggest takeaway from their work and research wasn't the 5-man infield they deployed on several occasions, but how objective they were in their decision making and that all it takes is a team executive willing to adopt unorthodox ideologies to achieve success in what is often seen as an unfair game.

Initial Baseball Background

Baseball has been at the forefront of using advanced metrics for player evaluation (Mondello, Kamke, 2014). This study's model was molded after the Weighted On-Base Average statistic. Weighted On-Base Average, created by baseball analyst Tom Tango, is a metric used to evaluate a hitter's offensive value based on the relative values of each distinct offensive event (Fangraphs).

One of the main reasons why Weighted On-Base Average was selected to be replicated was due to the importance it serves in baseball and the frequency with which it is used. The metric serves multiple purposes; Weighted On-Base Average alone has merit, and it can be easily converted into Weighted Runs Above Average (Fangraphs). Another reason for modeling Weighted Offensive Productivity Rating after Weighted On-Base Average, is the relatively simple concept behind the metric. The measure combines all aspects of hitting into one metric based on their proportion to their run value (Fangraphs).

From the previously mentioned War-on-Ice website, Thomas (2014) outlines the benefits derived from previous the attempts to develop single value metrics, and how these past attempts offer useful information for future models of evaluating player performance. The biggest takeaways from the article include:

1. Metrics that account for teammates and opponent quality are better than metrics that don't.
2. Incorporating various event types in models perform better than models that focus on fewer events or aspects of the sport.

3. An ideal model avoids using constants for the sake of using constants during calculations without offering valid justifications.

The proposed Weighted Offensive Productivity Rating satisfies the two latter and the importance of the former is tested in this case at the game level.

Translating Baseball Ideas to Hockey

It is frequently noted, in any attempt to use advanced metrics to value hockey players, that hockey is not baseball (Macdonald, 2012). While this is commonly made in jest or as a rhetorical statement, the idea behind it is that baseball is a series of independent situations with finite beginnings and endings to each situation, while hockey is a continuously flowing game with interdependent, unclear events. The process of transforming Weighted On-Base Average into Weighted Offensive Productivity Rating, focuses primarily on determining what a hitter's outcomes would look like for a hockey player. The process started with generalizing a hockey player's skillset into different components that met two criteria:

1. The components needed to be quantifiable
2. The components needed to be measured appropriately and accurately

Conversations with industry professionals and prior knowledge of the sport were the primary sources of information for developing the components of hockey's version of Weighted On-Base Average. It was decided that a hockey players' offensive productivity would be composed of his passing, zone transition, and shooting abilities. In terms of measuring these components, prior research and previous findings proved to be most useful. Stimson (2006) exhibited that shot assists are a repeatable measure, in addition to being predictive of future goal scoring. These two takeaways were the key factors in accepting primary and secondary shot assists as the measures for the model's passing component.

While identifying shooting talent or "finishing ability" is a difficult task given the current state of hockey analytics, Perry (2016) developed an expected goals model accounting for factors such as shot distance from the net, shot angle "in absolute degrees from the central line normal to the goal line", shot type, and strength state. Through his analysis, it was determined that taking shot attempts from short distance allows us to reasonably expect to score more goals as evidence in Figure 2, thus deeming shot attempts based on their "danger" (essentially distance and angle) from Figure 1 an appropriate measure of a player's shot quality.

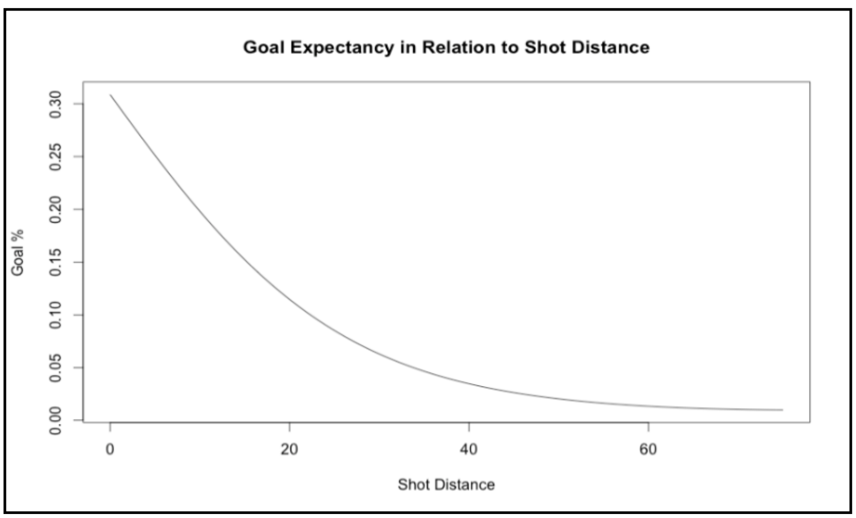

Figure 2. Goal Expectancy in Relation to Shot Distance. Reprinted from Corsica, by Perry.E, 2016, Retrieved from corsica.hockey.

One of the first in-depth statistical analyses of neutral zone play and zonal transitions was done by Tulsky et al. (2015), where the authors manually tracked 330 games from the 2011-2012 NHL season and found one of more impactful discoveries in hockey analytics in recent years. Their results showed that teams are able to generate roughly twice as many shot attempts and goals via controlled zone entries. From the other perspective, teams also surrendered fewer shot attempts and goals when they forced the opposition to "dump and chase" on their zone entries. These significant results warranted controlled zone entries and controlled zone exits to serve as the measures for the zone transition component of the Weighted Offensive Productivity Rating.

Lastly, the proportional weights are currently calculated based on the number of a certain outcome that lead to a goal divided by the total amount of that certain outcome. While it is not calculated the exact same way as Weighted On-Base average, Fangraphs calculates the linear weights using a run expectancy matrix based on all possible "base-out states" for each component of Weighted On-Base Average, it is in theory comparable.

Model Techniques

Throughout the research process, there were numerous different types of research and testing methods used. The studies composed of techniques deemed most applicable included:

1. Logistic regression (Baumer et al, 2015)
2. Split half reliability (Tulsky et al, 2013)
3. Cluster analysis (Vincent, Eastman, 2009)

Baumer et al. (2015) offered two key contributions to this study; the application of logistic regression to determining the value of outcomes and the concept of offering reproducibility to the study. A logistic regression model was fit to determine the probability of a baseball fielder recording an out on a ball in play based on x,y coordinates of the field displayed in Figure 3. Also, a significant portion of the study emphasizes the idea of providing readers with all the information needed in order to reproduce the exact study, allowing them to make any appropriate modifications that the reader might deem advantageous to the model. That concept is important in any statistical analysis, as the idea of “black-box” analysis frequently results in skepticism.

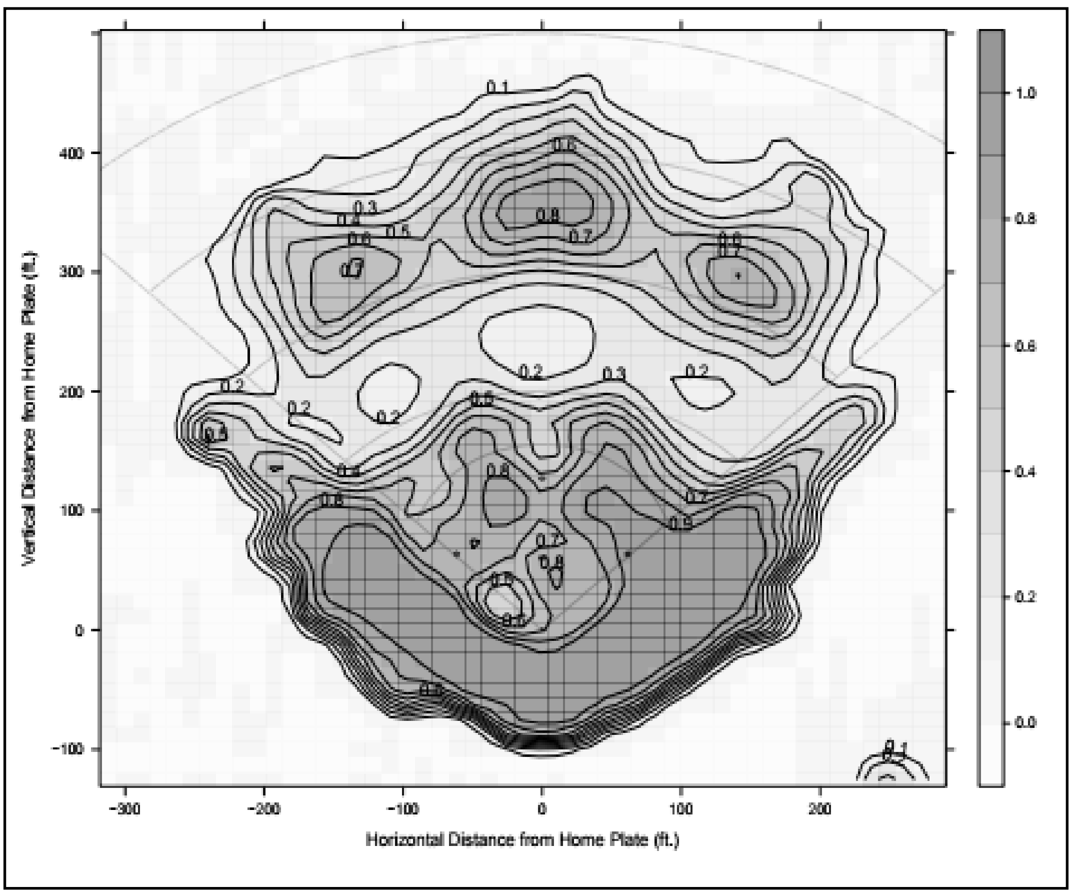

Figure 3. OpenWAR out probability plot. Reprinted from "OpenWAR: An open source system for evaluating overall player performance in major league baseball," by Baumer, B. S. et al., 2015, *Journal of Quantitative Analysis in Sports*, *11*(2), 69-84. Copyright [2015] of Journal of Quantitative Analysis in Sports.

Additionally, a split half reliability test may be appropriate to determine the reliability of the model from player to player and from game to game in the way Tulsky et al. (2013) tested their Zone Performance Scores. Testing the correlation between winning odd and even number games using both Weighted Offensive Productivity Rating and Points would appear to be most applicable.

Lastly, a K-means cluster analysis may be an interesting technique to use, not necessarily in the model itself, but once the model is complete to partition players into certain player types based on the passing, transition, and shooting model components (Vincent, Eastman, 2009). This analysis wouldn't necessarily factor into the model's predictive power of game outcomes, as it would be more descriptive in nature as opposed to predictive.

Chapter 3

Methodology

This chapter will focus on the procedures used for data collection and analysis in this study. The first section will further describe the subjects used in this study, and how these subjects were selected. This is followed by the description of the research design. The next section will explain the instruments and tools used by the study to collect and organize the data set. The procedures for data collection follows, where the process will be outlined in a manner in which it can be reproduced. Lastly, the data management and analysis section concludes this chapter, which includes the specific statistical analyses and techniques used in the study and why these tests were deemed appropriate.

Subjects

The subjects of this study are the members of Robert Morris University Men's Division I Ice Hockey team. This study measured the offensive abilities of only the team's forwards and defensemen, while goalies were excluded entirely from this study. To qualify for this study, a player will have played in at least 50% of the team's games from this season. These requirements give us a total of 23 eligible players of the 23 skaters on the roster in this case, consisting of 15 forwards and 8 defensemen.

Research Design

The research design of this study can be described as a descriptive correlational design. Data is collected over the course of the season and without making any changes to the study's environment or subjects (Grimes, Schulz, 2002). This descriptive correlational design provides a strong fit to test the research hypotheses for this study, as over the course of the season, the study was able to determine if Weighted Offensive Productivity Rating is correlated with player and team success, as well as alternative metrics. A logistic regression model was used, with the dependent variable being whether or not the team won or lost each game that each player played in. The likelihood of a win (dependent variable) was estimated as a function of several independent variables including:

1. Weighted Offensive Productivity Rating for player *i* in game *j* (wOPR*ij*)

2. Weighted Offensive Productivity Rating for all players not equal to player *i* in game *j* (TeamwOPR*ij*)
3. Total points for player *i* in game *j* ($Points_{ij}$)
4. Total points for all players not equal to player *i* ($TeamPoints_{N \neq i}$)
5. Winning percentage of opposing team (OppWP)
6. A dummy variable indicating a home or away game location (Home)

Quality of Competition was defined as that opposition's winning percentage for the season. NCAA hockey teams are slightly different than NHL teams for example, in the sense that they are unable to trade for players or sign free agents. When the NCAA season starts, the players on roster are largely going to be the players on roster for the final game of the season. Due to this, it was concluded that the team's winning percentage would be an acceptable measure of the opposition's abilities, regardless of time. In the future, a better measure of opponent quality across levels might be team winning percentage entering the game, or the Weighted Offensive Productivity Rating for that team.

Instrumentation

Due to the NCAA not collecting the specific data required for this study, the data used in this study are primary data collected from analyzing game film (Thomas, 2006). The game film is recorded live during each game using a video camera and DVSport software. The data is collected and organized using Microsoft Excel 2016. Two identified threats to internal validity include maturation and instrumentation. It is acknowledged that at this age, over the course of time, players will mature physically. Additionally, having multiple data trackers opens the possibility to scorer variability during the data collection process (Tulsky, et al., 2011). Inter-coder reliability was tested through the primary researcher “blindly” coding a randomly selecting a game coded by each data tracker. The primary researcher had an immaterial difference to the data trackers results (as shown in Appendix C).

The performance measure used in this study is the Weighted Offensive Productivity Rating (wOPR). The formula for the overall model of the metric is:

$$wOPR = Passing\ Component + Shooting\ Component + Zone\ Transition\ Component$$

Where:

$$Passing\ Component = (A2_{wA2} + LSA1_{wLSA1} + MSA1_{wMSA1} + HSA1_{wHSA1})$$

$$Zone\ Transition\ Component = (CE_{wCE} + CX_{wCX})$$

$$Shooting\ Component = (LSA_{wLSA} + MSA_{wMSA} + HSA_{wHSA})$$

And:

$$w_x = \frac{Number\ of\ x\ events\ that\ lead\ to\ a\ goal}{All\ x\ events}$$

Procedures

The data used in this study was collected over the course of the 2016-2017 NCAA Men’s ice hockey regular season. This process occurred from October of 2016 through April of 2017. During this time frame, the study used a rotating group of 4 data trackers. The data collection process was mapped out prior to the start of the season to ensure a standardized procedure to ensure data trackers had set guidelines throughout the season. Data trackers were instructed to analyze game film while simultaneously recording data points in Microsoft Excel 2016 workbooks issued by the team's Hockey Analytics Consultant. Each row in the Excel sheet was to contain an instance or event that would then be detailed by attributes in the following columns. For example, the event in row 1 is a shot attempt. The columns will include information such as the opponent, the player shooting the puck, where on the ice the shot occurred, the player(s) assisting on the shot attempt, and whether or not the shot resulted in a goal. Upon the completing of each game, the data trackers would forward the collected data points to the Hockey Analytics Consultant. This individual would then compile and organize the data in an Excel "master file" that was used for following data analysis.

Data Analysis

The data used in this study was collected and organized using Microsoft Excel 2016 workbooks, which were then loaded into Stata IC v14, where logistic regression was used to determine the likelihood of earning a win. In this study, the game outcome was coded using a binary variable where 1 equals a win and a 0 equals a loss.

This study used logistic regression to predict a binary dependent variable based on a combination of several independent variables. Key assumptions of logistic regression include:

1. Independence of observations
2. Little to no multicollinearity among independent variables
3. Independent variable to case ratio of at least 1 to 10.

In order to determine if Weighted Offensive Productivity is a valid indicator of offensive performance, this study tested the fit and performance of two separate logistic regression models. The first model will determine whether the Weighted Offensive Productivity both at the individual and team level are significant predictors of the likelihood of winning a game, controlling for opponent quality and home ice. The second model instead used Points, individual and team, as predictors of the likelihood of a team win using identical controls.

Table 1

Logistic Regression Variables

Variable Name	Variable Definition
wOPR	Weighted Offensive Productivity Rating for player *i* in game *j*
TeamwOPR	Weighted Offensive Productivity Rating for all players not equal to player *i* in game *j*
Points	Total points for player *i* in game *j*
TeamPoints	Total points for all players not equal to player *i*
OppWP	Winning percentage of opposing team
Home game	A dummy variable indicating a home or away game location

Two estimation methods were employed to guard against likely non-independence of game observations throughout the season. Standard, and player fixed-effects panel, logit estimations were generated and the resulting coefficients were compared. The formal model was as follows:

$$logit(win) = \beta_0 + \beta_1 wOPR + \beta_2 TeamwOPR + \beta_3 OppWP + \beta_4 Home + \varepsilon$$

And:

$$logit(win) = \beta_0 + \beta_1 Points + \beta_2 TeamPoints + \beta_3 OppWP + \beta_4 Home + \varepsilon$$

The results of these models were evaluated based on the pseudo R^2 that each model produces. The model that produced the higher pseudo R^2, may be considered the better indicator of on-ice performance. In a more granular sense, the first two research hypotheses were evaluated based on the t-statistic associated with the Beta coefficients for Weighted Offensive Productivity Rating and Points at the individual and team level.

The last hypothesis regarding small dataset situations was tested based on how the models perform during the first 10 games of the season, given the fact that the team only plays 34 games. In order to evaluate the performance of the two predictors, wOPR v. PTs, a prediction equation was constructed from the estimated parameters, and the player current pre-game metric (season to date average) were used to generate game predictions (expected win probabilities) for each game in the season. These predictions were then compared to actual game results a success/failure basis, using $p \geq 0.5$ as a win, and $p < 0.5$ as a predicted loss.

In order to illustrate the game by game updating of information, the cumulative average metrics were calculated for both the wOPR and Points, e.g. "season to date" per game wOPR and per game Points before game *j*. The ratio of these metrics to their season long average was then calculated to examine the relative variability of current performance metric information at each game of the season, but also the comparative deviation between wOPR and Points as the season progresses.

As a further analysis to test the relationship between the two evaluation metrics being tested, a Pearson's Correlation and Spearman's Rank Correlation was generated for the players' wOPR per game and Points per game. This served as an initial test of the wOPR to determine if the metric was statistically related to a previously accepted performance metric such as Points per game over a large number of games (e.g. 30+).

Chapter 4

Results

The four model estimation procedures generated parameter estimates with statistically significant model fit ($\alpha = 0.05$), in addition to logistic regression pseudo R-squared values of 0.068 and 0.146 for wOPR and Points, respectively. The beta coefficient estimates across the pooled data and fixed-effects panel logit were very consistent. It appeared that any player fixed-effects on the game outcomes were minimal, independent of the performance metrics that were included in the models. In the wOPR models, team wOPR and the opponent's winning percentage both had statistically significant relationships to the game outcome in the expected directions, positive and negative respectively in Table 2. In short, as a team increases their wOPR or faces an opponent of lower quality, they are more likely to win that game. The Points models identified the same relationships, but further identified a positive home ice advantage and that individual points was significant at the $\alpha = 0.1$ level, as displayed in Table 3.

Table 2 *wOPR Logistic Regression Table*				
Variable Name	Logit		Panel Logit	
	Beta	P-value	Beta	P-value
wOPR	0.167	0.609	0.196	0.632
TeamwOPR	0.214	<0.001*	0.203	<0.001*
OppWP	-4.854	<0.001*	-4.713	<0.001*
Home game	0.014	0.935	0.015	0.927
intercept	1.070	0.040	n/a	n/a
Notes. Pseudo R^2 = 0.068 (*p < .05)				

Table 3 *Points Logistic Regression Table*				
Variable Name	Logit		Panel Logit	
	Beta	P-value	Beta	P-value
Points	0.315	0.071	0.317	0.82
TeamPoints	0.325	<0.001*	0.311	<0.001*
OppWP	-3.709	<0.001*	-3.620	<0.001*
Home game	0.833	<0.001*	0.803	<0.001*
intercept	0.152	0.742	n/a	n/a
Notes. Pseudo R^2 = 0.146 (*p < .01)				

In order to illustrate the game by game updating of information, the cumulative average metrics were formulated for both the wOPR and Points, e.g. “season to date” per game wOPR before game *j*. The simple ratio of these metrics to their season long average was calculated to examine the relative variability of current performance metric information. The graphs in Figure 4 and Figure 5, which have identical y-axis scaling, both illustrate the classic statistical regression phenomena, but also highlight the comparative deviation between wOPR and Points early in the season.

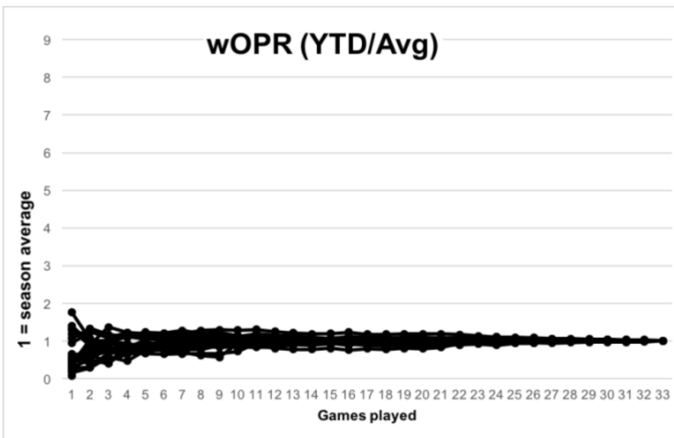

Figure 4. Ratio of year-to-date wOPR to season average wOPR.

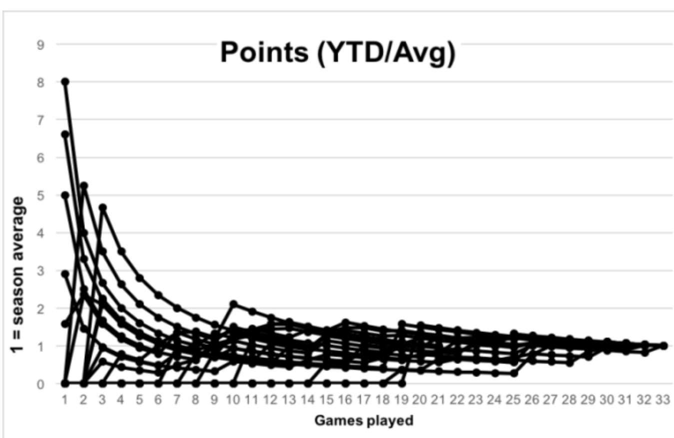

Figure 5. Ratio of year-to-date Points to season average Points.

The results of the Pearson's R correlation and Spearman's rank correlation are also displayed in Figure 7. Additional qualitative differences between the metrics are displayed in Figure 6, by contrasting the ability of the following prediction equations generated from the logistic regression procedures to identify game outcomes.

$$P(win) = \frac{e^{\beta 0} + \beta 1wOPR + \beta 2TeamwOPR + \beta 3OppWP + \beta 4Home}{1 + e^{\beta 1wOPR + \beta 2TeamwOPR + \beta 3OppWP + \beta 4Home}}$$

And:

$$P(\mathrm{win}) = \frac{\mathrm{e}^{\beta 0} + \beta 1\mathrm{Points} + \beta 2\mathrm{TeamPoints} + \beta 3\mathrm{OppWP} + \beta 4\mathrm{Home}}{1 + \mathrm{e}^{\beta 1\mathrm{Points} + \beta 2\mathrm{TeamPoints} + \beta 3\mathrm{OppWP} + \beta 4\mathrm{Home}}}$$

Finally, Table 4 shows that the wOPR correctly predicted 20 games over the course of the year, and the Points model predicted a similar 21 of the 34 games correctly.

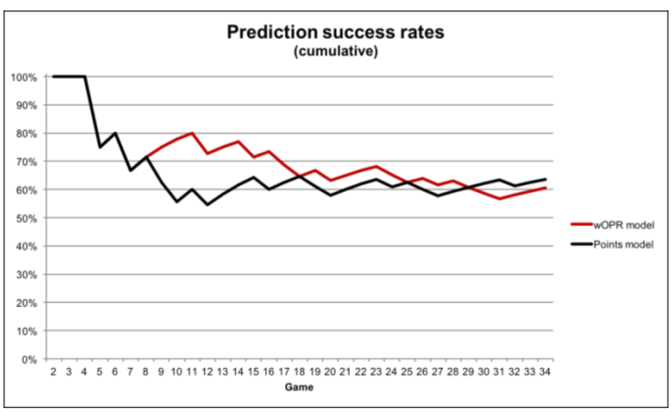

Figure 6. Prediction success rate of wOPR and Points prediction models.

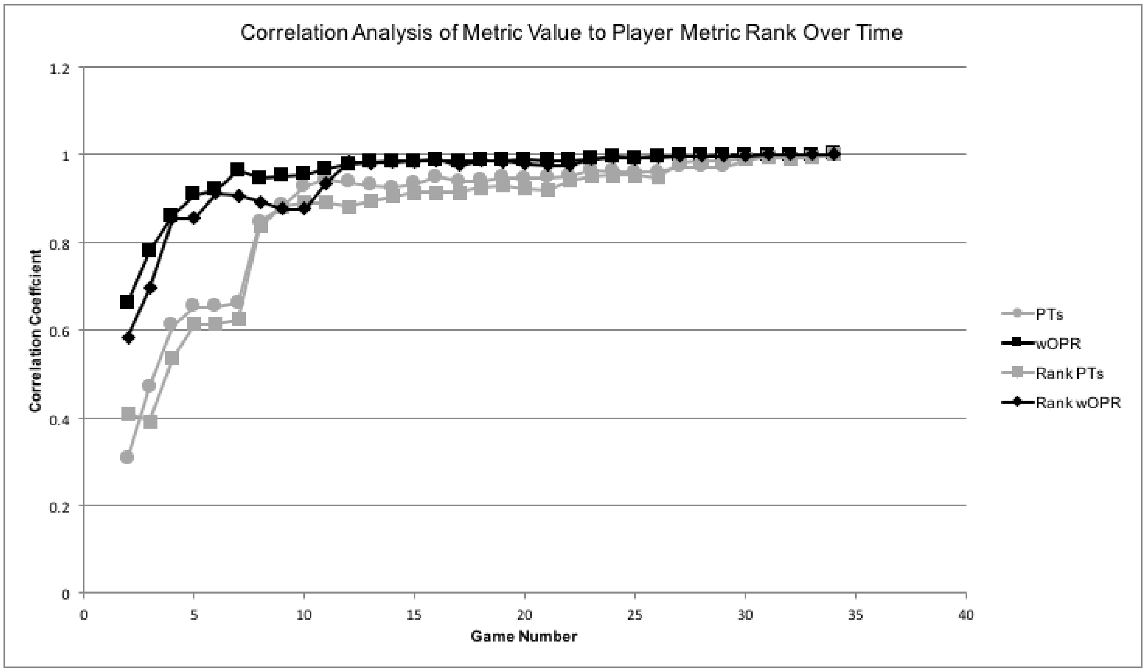

Figure 7. Correlation of average wOPR and Points per game to player rank in wOPR and Points per game.

Game	Pr(Win)OPR	Pr(Win)PTS	Outcome
2	0.526	0.607	1
3	0.294	0.264	0
4	0.326	0.292	0
5	0.410	0.311	1
6	0.409	0.288	0
7	0.342	0.202	1
8	0.352	0.239	0
9	0.587	0.433	1
10	0.614	0.444	1
11	0.669	0.552	1
12	0.643	0.551	0
13	0.628	0.503	1
14	0.623	0.501	1
15	0.602	0.474	0
16	0.616	0.486	1
17	0.605	0.479	0
18	0.590	0.467	0
19	0.616	0.478	1
20	0.394	0.264	1
21	0.696	0.569	1
22	0.694	0.572	1
23	0.278	0.196	0
24	0.299	0.198	1
25	0.544	0.404	0
26	0.537	0.409	1
27	0.628	0.526	0
28	0.633	0.517	1
29	0.622	0.497	0
30	0.596	0.462	0
31	0.603	0.463	0
32	0.580	0.454	1
33	0.835	0.753	1
34	0.842	0.770	1

Figure 8. Results of wOPR and Points predictions models per game.

Based on this information, we would fail to reject the null hypotheses of wOPR showing no relationship to on-ice performance at the individual player level. However, we would reject the null hypotheses regarding wOPR at the team level and in small dataset situations. The results of the logistic regression models showed that, while in the positive direction, wOPR at the individual player level was not statistically significant at the $\alpha = 0.05$ level. The wOPR at the team level however proved to be statistically significant at the $\alpha = 0.05$ level and the data in Figure 4 displayed that wOPR is less volatile while, as an example, offering a value for a player in game 4 much more comparable to the value in game 34 compared to Points, which shows much more variability over the course of the season.

Chapter 5

Discussion

The problem many teams have is a lack of accurate performance metrics to objectively evaluate their team and individual players on a short-term basis. The above results indicate that wOPR was a statistically significant predictor of performance at the team level. In application, increasing a team's wOPR should increase their likelihood of winning. Shooting often, and from high danger areas, often results in an increase in goal scoring and is predictive of future goal scoring (Perry, 2016). While both metrics performed well late in the season, wOPR appeared to perform adequately (R > 0.70) earlier in the season (Figure 8) and displayed better predictive performance through the middle of the season (Figure 6), after which the Points metric performed similarly, if not slightly better.

The wOPR at the individual player level did not prove to be a statistically significant predictor of team success in this study. However, a larger data set might have improved the statistical power to identify what are likely small individual effects. While it is difficult to know for sure if this will result in a statistically significant predictor variable, a data set consisting of several teams and/or several years is certainly something to consider in future studies. We already know that controlled zone entries lead to more goals scored, and shot assists are a repeatable measure predictive of future goal scoring (Tulsky, et al., 2013; Stimson, 2016). One could presume that more data, and consequentially more refined weights, might produce results more favorable for the wOPR model. While the weights used in the study are appropriate and calculated in a manner that is analogous in theory to the way that wOBA calculates the linear weights in its formula, the small dataset situation at hand makes the linear weights used in wOPR a potential limitation to the study. A solution to this problem would be collecting and analyzing data from several teams over several seasons in order to refine the linear weights, however as previously mentioned, data at the NCAA level is suboptimal and would require an extensive amount of time and manpower.

Lastly, it would be interesting to dive into the game film and look some of the underlying numbers with regards to the single game predictions. It might be thought-provoking to look at the games the wOPR model predicted incorrectly and look to see why this might be the case. One initial factor that currently isn't considered by the model is power-play and penalty-kill situations. This particular team had the 7th best power-play in the nation, clicking at a 22.3% rate (NCAA, 2017)**.** As advantageous as additional teams or seasons of data may be, accounting for a team's ability to produce with the man advantage could be equally as important to improving wOPR's predictive power. On a similar note, in addition to accounting for special team's play, there could be several aspects that could improve the prediction models including: a more refined quality of competition measure, such as the opposition's wOPR; a measure for both team's goaltending abilities; and a measure of each team's ability to play defense.

For a hockey team or organization looking to implement this model into their decision-making process, the application and requirements to do so would be manageable to teams unable to afford the luxury of having their own team of quantitative analysts on staff.

While most hockey teams, especially those with limited data analytics capabilities, are lightyears away from making player personnel solely off of the results of a statistical model, a coach could be in the position to give his team a competitive advantage over the opposition were he to decide to include players in the lineup with higher wOPR. Often times it can be difficult to differentiate between the abilities of players using subjective qualitative analysis or basic counting statistics. This gap can be filled by the wOPR model.

A team could begin to track wOPR for the players on their team using game film, basic competency in Microsoft Excel, and a crash course in statistics. The process could be completed by a member of the coaching staff, a small group of volunteers, or a combination of the two.

Conclusion

The wOPR model provides decision makers with an additional tool to use when making lineup decisions from the short-term point of view. wOPR at the team level proved to be a statistically significant predictor of future team success, as the wOPR model (formulated with only 5 on 5 play data), predicted roughly 60% of the 34 games from the 2016-2017 season correctly for this NCAA Men's Division I hockey team. While the wOPR model was not as strong of a predictor of team success as total points, it was a better indicator of underlying player ability in the short term when points per game rates fluctuate heavily and coaches have limited data on new players.

References

Analytics. (n.d.). *Dictionary.com Unabridged*. Retrieved January 11, 2018 from Dictionary.com website http://www.dictionary.com/browse/analytics

Assumptions of Logistic Regression. (n.d.). In Statistics Solutions. Retrieved January 8, 2018, from http://www.statisticssolutions.com/assumptions-of-logistic-regression/

Baumer, B. S., Jensen, S. T., & Matthews, G. J. (2015). OpenWAR: An open source system for evaluating overall player performance in major league baseball. *Journal of Quantitative Analysis in Sports*, *11*(2), 69-84. Retrieved from EBSCOhost.

Grimes, D. A., & Schulz, K. F. (2002, January 12). Descriptive studies: what they can and cannot do [Electronic version]. *The Lancet*, *359*(9301), 145-149.

Lewis, M. (2004). *Moneyball*. New York, NY: W. W. Norton Company, Inc.

Lindbergh, B., & Miller, S. (2016). *The Only Rule Is It Has to Work*. New York, NY: Henry Holt and Company, LLC.

Linear Weights. (n.d.). In *Fangraphs*. Retrieved from https://www.fangraphs.com/library/principles/linear-weights/

Macdonald, B. (2011). A Regression-Based Adjusted Plus-Minus Statistic for NHL Players. *Journal of Quantitative Analysis in Sports*, *7(3).* Retrieved from De Gruyter.

Macdonald, B. (2012). Adjusted Plus-Minus for NHL Players using Ridge Regression with Goals, Shots, Fenwick, and Corsi. *Journal of Quantitative Analysis in Sports*. Retrieved from De Gruyter.

Marek, P., Sediva, B., & Toupal, T. (2014). Modeling and prediction of ice hockey match results. *Journal of Quantitative Analysis in Sports*, *10*(3). Retrieved from De Gruyter.

Men's Ice Hockey Statistics-Team. (2017, April 11). In *NCAA*. Retrieved from http://www.ncaa.com/stats/icehockey-men/d1/current/team/182

Mondello, M., & Kamke, C. (2014). The Introduction and Application of Sports Analytics in Professional Sport Organizations: A Case Study of the Tampa Bay Lightning. *Journal of Applied Sport Management*, *6*(2). Retrieved from EBSCOhost.

Neal, D., Tan, J., Hao, F., & Wu, S. S. (2010). Simply Better: Using Regression Models to Estimate Major League Batting Averages. *Journal of Quantitative Analysis in Sports*, *6*(3). Retrieved from De Gruyter.

Perry, E. (2016, August 13). Shot Quality and Expected Goals: Part 1.5. In *Corsica*. Retrieved from http://www.corsica.hockey/blog/2016/08/13/shot-quality-and-expected-goals-part-1-5/

Perry, M. (2017, May 20). The Art of WAR. In *Corsica*. Retrieved June 16, 2017, from http://www.corsica.hockey/blog/2017/05/20/the-art-of-war/

Schuckers, M., & Curro, J. (2013). Total Hockey Rating (THoR): A comprehensive statistical rating of National Hockey League forwards and defensemen based upon all on-ice events. In *MIT Sports Analytics Conference*. Retrieved December 19, 2016, from http://www.sloansportsconference.com/wp-content/uploads/2013/Total%20Hockey%20Rating%20(THoR)%20A%20comprehensive%20statistical%20rating%20of%20National%20Hockey%20League%20forwards%20and%20defensemen

Stimson, R. (2016, January 26). Redefining Shot Quality: One Pass at a Time. In *Hockey-Graphs*. Retrieved from https://hockey-graphs.com/2016/01/27/redefining-shot-quality-one-pass-at-a-time/

Thomas, A. C. (2006). The Impact of Puck Possession and Location on Ice Hockey Strategy. *Journal of Quantitative Analysis in Sports, 2(1)*. Retrieved from De Gruyter.

Thomas, A. C. (2014). The Road to WAR. In *War On Ice*. Retrieved November 11, 2016, from http://blog.war-on-ice.com/index.html%3Fp=429.html

Thomas, A. C. (2014, October 5). The Road to WAR (for hockey), Part 1: The Single-Number Dream. In War On Ice: The Blog. Retrieved January 8, 2018, from http://blog.war-on-ice.com/index.html%3Fp=37.html#note-1

Tulsky, E., Detweiler, G., Spencer, R., & Sznajder, C. (2013). Using Zone Entry Data to Separate Offensive, Neutral, and Defensive Zone Performance. In *MIT Sloan Sport Analytics Conference*. Retrieved February 20, 2017, from http://www.sloansportsconference.com/wp content/uploads/2013/Using%20Zone%20Entry%20Data%20To%20Separate%20Offensive,%20Neutral,%20And%20Defensive%20Zone%20Performance.pdf

Vincent, C. B., & Eastman, B. (2009). Defining the Style of Play in the NHL: An Application of Cluster Analysis. *Journal of Quantitative Analysis in Sports*, *5*(1). Retrieved from De Gruyter.

Wheeldon, J. (2017, September 25). The Future of Hockey Analytics. In *The Hockey Writers*. Retrieved January 11, 2018, from https://thehockeywriters.com/future-of-hockey-analytics/

wOBA. (n.d.). In *Fangraphs*. Retrieved from http://www.fangraphs.com/library/offense/woba/

Appendix A: Literature Review Search Results

Search Engine	Search Terms	# Retrieved	# Met Inclusion Criteria	Table 2 ID#
De Gruyter	hockey	97	5	1,2,7,8,10
De Gruyter	WAR	21	1	3
De Gruyter	Weighted On-Base Average	17	1	5
Totals		135	8	
Google	Tulsky sloan research paper	34,600	1	6
Google	Weighted On-Base Average	8,520	1	9
Google	Logistic Regression Assumptions	393,000	1	18
Google	wOBA linear weights	10,400	1	17
Google	RMU men’s hockey statistics	50,700	1	16
Google	Hockey expected goals	14,300,000	1	15
Google	Hockey WAR	38,000,000	3	12,14,19
Google	Hockey shot quality	9,880,000	1	13
Google	Schuckers sloan research paper	12,600	1	20
Google	Analytics definition	5,010,000	1	21
Google	Hockey analytics debate	1,580,000	1	22
Total		69,279,820	12	
Google Scholar	Hockey performance evaluation metrics	19,200	0	
Google Scholar	Hockey performance metrics	18,600	0	
Totals		37,800	0	
SportDiscus	Sport analytics	139	1	4
Totals		139	1	

Appendix B: Literature Review Table

	Author(Year)	Study Design	Sample Size	Key Findings	Include/Exclude
1	Macdonald, B. (2012). "Adjusted Plus-Minus for NHL Players using Ridge Regression with Goals, Shots, Fenwick, and Corsi." Journal of Quantitative Analysis in Sport Manuscript 1447.	Ridge Regression	N = 2,324,528 observations	Ridge estimates of shots, Fenwick, and Corsi more consistent than goals. Even-strength estimates offer higher year-to-year correlations than PP and PK estimates	Include
2	Thomas, A. C. (2006). "The Impact of Puck Possession and Location on Ice Hockey Strategy." Journal of Quantitative Analysis in Sports **2**(1).	Semi-Markov Process	N= 18 Games from 2004-2005 NCAA Hockey season	Teams executing a "carry-in" offense are more likely to score. Teams executing a "dump-chase" offense are less likely to be scored on	Include
3	Baumer, B. S., S. T. Jensen, et al. (2015). "openWAR: An open source system for evaluating overall player performance in major league baseball." Journal of Quantitative Analysis in Sports **11**(2): 69-84.	Logistic Regression	N= 1284 players from 2012 MLB season N= 1303 players from 2013 MLB season	Within-player autocorrelation of openWAR shared similar results to rWAR and fWAR	Include
4	Mondello, M. and C. Kamke (2014). "The Introduction and Application of Sports Analytics in Professional Sport Organizations." Journal of Applied Sport Management **6**(2).	Definitions and theory	N/A		Include
5	Neal D., J. Tan, et al. (2010). "Simply Better: Using Regression Models to Estimate Major League Batting Averages." Journal of Quantitative Analysis in Sports **6**(3).	Linear Regression	N= Undefined	Number of at-bats vital predictor of batting average. Linear regression models proved powerful compared to Bayesian estimators	Exclude
6	Tulsky, E., G. Detweiler et al. (2011). "Using Zone Entry Data To Separate Offensive, Neutral, and Defensive Zone Performance."	Split-half reliability test	N= 330 games from 2011-2012 NHL season	Controlled entries result in more shot attempts and goals. Failed controlled entries lead to fewer shots against than "dump-chase" plays.	Include
7	Macdonald, B. (2011). "A Regression-Based Adjusted Plus-Minus Statistic for NHL Players." Journal of Quantitative Analysis in Sports **7**(3).	OLS Regression	N= 798,214 shifts from 2007 – 2010 NHL seasons	Did not account for interaction between teammates. Acknowledged doing so could reduce error terms.	Include
8	Vincent, V. B. and B.	Cluster	N= 625 players	3 clusters (player	Include

	Eastman (2009). "Defining the Style of Play in the NHL: An Application of Cluster Analysis." Journal of Quantitative Analysis in Sports **5**(1).	Analysis	from 2002-2003 NHL season	types). Scorers were highest earners. No earnings difference between Grinders/Enforcers.	
9	"wOBA." Fangraphs.	Definitions	N/A		Include
10	Marek, P., B. Sediva, et al. (2014). "Modeling and prediction of ice hockey match results." Journal of Quantitative Analysis in Sports **10**(3): 357-365.	Poisson Distribution	N= 3193 games from 1999-2010 Extraliga seasons N= 363 games from 2011 Extraliga season	Predictive ability of the models offers potential for profit to be made against bookmakers.	Include
11	Grimes, MD, D.A. and K. F. Schulz (2002)."Descriptive studies: what they can and cannot do." The Lancet **359**(9301).	Theory and definitions	N/A		Include
12	Thomas, A. C. (2014, October 5). "The Road to WAR (for hockey), Part 1: The Single-Number Dream." War On Ice: The Blog.	Theory and definitions	N/A		Include
13	Stimson, R. (2016, January 26). "Redefining Shot Quality: One Pass at a Time." Hockey-Graphs.	Theory and definitions	N/A		Include
14	Perry, M. (2017, May 20). "The Art of WAR." Corsica.	Theory and definitions	N/A		Include
15	Perry, E. (2016, August 13). "Shot Quality and Expected Goals: Part 1.5." Corsica.	Theory and definitions	N/A		Include
16	"Men's Ice Hockey Statistics-Team. (2017, April 11)." NCAA	Public Data Source	N/A		Include
17	"Linear Weights." Fangraphs	Theory and definitions	N/A		Include
18	"Assumptions of Logistic Regression." Statistics Solutions	Theory and definitions	N/A		Include
19	Thomas, A. C. (2014). "The Road to WAR." War On Ice: The Blog	Theory and definitions	N/A		Include
20	Schuckers, M., & Curro, J. (2013). "Total Hockey Rating (THoR): A comprehensive statistical rating of National Hockey League forwards and defensemen based upon all on-ice events." MIT Sports Analytics Conference.	Theory	N/A		Include
21	"Analytics." Dictionary.c	Theory and	N/A		Include

	om Unabridged	definitions			
22	Wheeldon, J. (2017, September 25). “The Future of Hockey Analytics.” The Hockey Writers.	Theory	N/A		Include
23	Lewis, M. (2004). “Moneyball.”	Theory	N/A		Include
24	Lindbergh, B., & Miller, S. (2016). “The Only Rule Is It Has to Work.”	Theory	N/A		Include

Appendix C: Inter-coder Reliability Results

Inter-coder Reliability	
	Percent of Agreement
Coder 1	88%
Coder 2	97%
Coder 3	93%