ANNUAL REVIEWS

Annual Review of Statistics and Its Application

Historical Perspectives and Current Directions in Hockey Analytics

Namita Nandakumar and Shane T. Jensen

Department of Statistics, The Wharton School, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA; email: namitan@wharton.upenn.edu, stjensen@wharton.upenn.edu

Annu. Rev. Stat. Appl. 2019. 6:19–36

First published as a Review in Advance on December 10, 2018

The *Annual Review of Statistics and Its Application* is online at statistics.annualreviews.org

https://doi.org/10.1146/annurev-statistics-030718-105202

Keywords

hockey, credit assignment, performance, strategy

Abstract

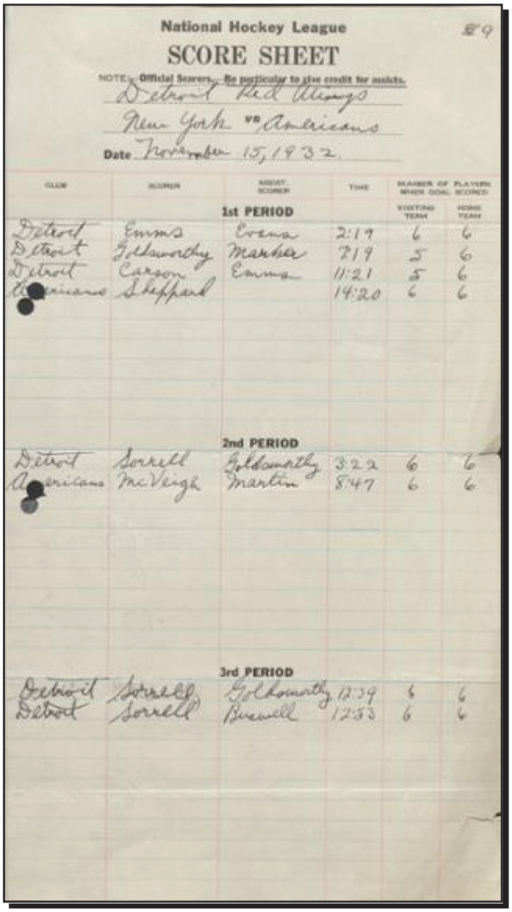
We review recent advances in hockey analytics research, most of which have occurred from the early 2000s to the present day. We discuss these advances in the context of earlier attempts to evaluate player performance in hockey. We survey the unique challenges of quantitatively summarizing the game of hockey, and how deficiencies in existing methods of evaluation shaped major avenues of research and the creation of new metrics. We present an extended analysis of the National Hockey League entry draft in terms of both retrospective evaluation and prospective strategy. We conclude with recommendations for future research in hockey analytics.

1. EARLY HISTORY OF HOCKEY ANALYTICS

Hockey is a fast, free-flowing game played on a sheet of ice, with six players per side attempting to score goals by using wooden sticks to propel a rubber puck into nets that are fixed at either end of the ice surface. Several aspects of the sport are a challenge for statistical evaluation, such as the continuity of the game play, the frequent substitution of individual players, and the infrequency of actual scoring events. In fact, there are fewer scoring events in hockey compared with any other major North American sport.

These inherent challenges in quantitatively summarizing the game of hockey means that we have not seen the kind of comprehensive *Moneyball* (Lewis 2004) revolution experienced by baseball and more recently basketball, where analytical tools have fundamentally changed every facet of the professional game, from roster construction to style of play.

That said, there are several directions of active research by which fans and analysts have sought to better understand how to understand and evaluate the sport of hockey, both on and off the ice. We can look to some of the oldest National Hockey League (NHL) box scores (**Figure 1**) to


National Hockey League

SCORE SHEET

NOTE.—Official Scorers.—Be particular to give credit for assists.

Detroit Red Wings

New York vs Americans

Date November 15, 1932.

CLUB	SCORER	ASSIST. SCORER	TIME	NUMBER OF PLAYERS WHEN GOAL SCORED: VISITING TEAM	HOME TEAM
		1st PERIOD			
Detroit	Emms	Evans	2:19	6	6
Detroit	Goldsworthy	Marker	7:19	5	6
Detroit	Carson	Emms	11:21	5	6
Americans	Sheppard		14:20	6	6
		2nd PERIOD			
Detroit	Sorrell	Goldsworthy	3:22	6	6
Americans	McVeigh	Martin	8:47	6	6
		3rd PERIOD			
Detroit	Sorrell	Goldsworthy	12:39	6	6
Detroit	Sorrell	Buswell	12:53	6	6

Figure 1

Score sheet from a National Hockey League game in 1932. Reproduced with permission from **https://www.nhl.com/news/nhl-historical-stats-now-available/c-291389804**.

illuminate some of the ways in which the field of hockey analytics has been uniquely shaped and guided by features of the sport.

We see in **Figure 1** that the summary statistics for individual player performance were restricted to tabulating which players scored goals and which players assisted in those goals. We see that these particular credits were not automatic in the early days of hockey, as evidenced by the note to "Be particular to give credit for assists."

Even when properly recorded, goals and assists are a limited summary for judging a player's performance as these scoring events are relatively infrequent during game play. In addition, not all players are directly tasked with scoring goals, such as defensemen and goaltenders.

An early attempt to reward players for goal prevention in addition to goals scored was the plus-minus statistic, which was officially tracked by the NHL starting in 1967–1968 (Desjardins 2009a). The plus-minus for each player is calculated as the goals scored by the player's team minus the goals scored against the player's team while that player is on the ice.

Plus-minus rewards players for positive contributions in terms of both goal scoring and goal prevention and so does not require a player to be offensively focused in order to perform well. It is relatively simple to tabulate and compare across players, which explains why it remains a popular metric up until the present day. However, plus-minus does suffer from some notable disadvantages.

One fundamental problem of plus-minus is the inability to separate an individual player's performance from the performance of their teammates (and opponents) who were also on the ice during each scoring event. Some players will have an undeservedly low plus-minus from playing on a bad team, whereas other players will disproportionately benefit from playing with teammates who are much better than themselves. In Section 3, we will review regression-based approaches to adjusting metrics such as plus-minus to account for teammates and opposition.

Another problem with plus-minus is that the relative infrequency of scoring events means that any metric based on goals provides a rather incomplete summary of all that happens during a typical hockey game. This problem is exacerbated by the fact that goals scored during power play situations are not counted toward plus-minus even though they are a substantial portion of the goals that are scored in hockey. The desire for a more complete representation of player performance based on more frequent game events has led to the more recent adoption of shot-based metrics, such as Corsi, which we discuss in more detail in Section 4 below.

2. MOVING BEYOND THE BOX SCORE

For decades, basic box score statistics such as plus-minus, combined with the eye test of scouts and coaches, were used to analyze teams and individual players in hockey. By the mid-2000s, however, the status quo began to change. The advent of Corsi, which we discuss in more detail in Section 4, is perhaps the most well-known development in the field of hockey analytics.

However, it is certainly not the only one. Gabe Desjardins's (2009f) advanced stats primer provides a nice summary of some of the first advances and priorities of hockey analysis. Its home on the Winnipeg Jets fan blog *Arctic Ice Hockey* alludes to how public hockey analytics advances have been driven largely by fan desire to better understand the sport.

Another direction of advancement was the development of metrics for evaluating quality of competition. By measuring the quality of all the opponents a player sees and weighting by shared ice time, we can identify players with varying defensive assignments and control for them (Desjardins 2009b). The degree to which quality of competition is significant, especially when taking quality of teammates into account, has been debated extensively (Desjardins 2009e). However, measuring the quality of teammates runs into the same difficulties as measuring a player's own performance, in that it is difficult to tease out the individual effects of players from performance of the team as a whole.

Another direction of work has been summarizing player usage through zone starts that quantify whether a player is deployed primarily in the offensive or defensive zone based on face-off counts (Desjardins 2009c). Zone starts can highlight differences between players in terms of their role in contributing to team success, but they are limited in terms of measuring the impact of each player in their role.

Penalty differential is another nonscoring related aspect of player performance that is relatively easily calculated based on box score statistics. The value in being able to put one's team on the power play more frequently has been recognized and has been demonstrated to be a repeatable skill in some players (Desjardins 2009d).

3. ADJUSTMENTS TO PLUS-MINUS

The well-known deficiency of plus-minus is that it tends to favor players on good teams and punish players on bad teams. From a statistical perspective, plus-minus is similar to a marginal effect in a regression model: The plus-minus for a particular player aggregates over all teammates and opposition who played with that player.

A more desirable metric of performance would be a partial plus-minus that adjusted for the teammates and opposition who played with that player. One simple adjustment for this is to subtract the off-ice plus-minus of their team from a player's on-ice plus-minus. This results in a statistic that has been termed relative plus-minus, which still does not tease out individual effects in cases of highly correlated player usage (Desjardins 2009a).

Regression models have also been employed as a more sophisticated approach to adjusting plus-minus to account for teammates and opponents. For example, Schuckers et al. (2011) analyzes player effects on all on-ice events, weighted by the change in probability of each team scoring within 10 seconds of the event. They ran a linear regression with indicator variables for each player on the ice and found that the resulting ratings were relatively uncorrelated with plus-minus and stable even when players changed teams, indicating that this metric effectively teased out individual player effects.

Gramacy et al. (2013) present a regularized logistic regression model for estimating true player plus-minus. The regularization component of their model was motivated by the persistent issue of multicollinearity when estimating player effects in hockey. Players tend to play as part of set lines that shift on and off the ice together, which means that the indicators of which players are on the ice during scoring events are highly collinear, which makes it very difficult to accurately estimate the partial effect of any one player.

Regularization of regression coefficients can help to stabilize estimates of the partial effects for each player. Gramacy et al. (2013) use a penalized regression method with team and individual player effects in order to identify players who have the most substantial impacts on goal scoring. They find that the adjusted partial effects on scoring are zero for a majority of players.

Thomas et al. (2013) also use regularization techniques to estimate player effects to identify the most consistently offensively and defensively impactful players. Their approach is based on modeling scoring in a hockey game as competing semi-Markov processes between the two teams based on the particular players on the ice at any given time. Over the five seasons of five versus five (5v5) player (i.e. both teams at even strength) situations analyzed, they find that only 2.3% of players had estimated player scoring effects that differed significantly from zero.

As noted by Thomas et al. (2013), these analyses are computationally expensive and mainly targeted to academic audiences. The result of the vast majority of player effects being shifted to zero is interesting but not overly helpful for the purposes of fans and analysts seeking to understand differences in player ability throughout a lineup.

Regression-based adjustments represent an important first step toward the larger goal of more holistic measures of player performance that give appropriate credit to the multiple ways in which a player can contribute to winning during game play. In Section 7, we examine recent efforts to emulate more comprehensive wins above replacement (WAR) metrics that have become commonplace in other sports, especially major league baseball.

4. SHOT-BASED MEASURES OF PERFORMANCE

As mentioned above, one of the greatest inherent challenges to statistical analysis of the sport of hockey is that scoring events are considerably rarer than in most other major sports. Thus, much of the effort in hockey analytics has been to find additional quantitative measures of game play beyond goals for and against. Given the infrequency of scoring, the primary focus of these efforts has been toward better measures of scoring opportunities, which leads most naturally to the tabulation of shots on goal.

Operating under the pseudonym Vic Ferrari, an Oilers blogger coined "Corsi" as the sum of shots on goal, missed shots, and blocked shots (McKenzie 2014). Corsi helps identify teams and individual players generating more scoring opportunities through shots, which ultimately should result in more goals and wins.

Several modifications of Corsi have sought to isolate the impact of individual players. Corsi relative to teammates compares the percentage of shot attempts for and against a team with and without a player on the ice. A related statistic, Fenwick, excludes blocked shots from the count. When a player blocks a shot, their Corsi differential is penalized, but their Fenwick differential is not. If the ability to block shots is repeatable and valuable, then Fenwick is the more appropriate metric to cite (Cane 2014).

One issue with shot-based metrics such as Corsi is that teams can alter their strategy toward taking more or fewer shots depending on the game situation. One example is that trailing teams shoot the puck more than leading teams as they attempt to make up the deficit, and this effect is amplified as the lead is extended, since the leading team has more of an incentive to play defensively.

These observations led to modifications of Corsi such as Corsi close, which restricts the calculation of shot attempts to situations where teams are within one goal of each other. Restricting the calculation of metrics based on game situation has been seen in other sports, such as American football, where performance in so-called garbage time situations is often downweighted in analyses (Salfino 2017).

However, the restriction of Corsi close to only game situations where there is a small goal differential does exclude a large amount of potentially useful data (Tulsky 2012). As an alternative approach, Micah Blake McCurdy (2014) developed a score-adjusted Corsi, which evaluates shot attempts relative to expectations based on the current score. Expectations that are also based on time remaining in the game were also considered but were determined to not be worth the additional complexity involved in accounting for time remaining. The score-adjusted Corsi metric was shown to have superior repeatability and predictive power over the course of a season over various other adjustments to Corsi and Fenwick (McCurdy 2014).

The increased emphasis on the repeatability and predictive power of metrics like Corsi and Fenwick has provided several insights into the relative roles of luck and skill in the game of hockey. One important finding is that for the vast majority of NHL players, percentage of shots converted to goals will regress to the overall means for their position (Tulsky 2013b). In a similar finding, secondary assists have been shown to be more statistically noisy than primary assists and goals scored, and they appear to be influenced more by randomness and team effects than a player's own skill (Tulsky 2011).

An ongoing criticism of Corsi and shot-based metrics in general is that these simple tabulations do not take into account the quality of each shot. Shots vary substantially in terms of their speed, location, and probability of scoring. Ideally, we would give extra reward to players who are able to take their shots in more optimal locations and game situations. In the next section, we review attempts to account for the quality of scoring opportunities.

5. MEASURES OF SHOT QUALITY AND EXPECTED GOALS

The earliest comprehensive quantitative study of the concept of shot quality is an article from Alan Ryder (2004a). This overview is also an interesting microcosm for several inherent problems with using NHL data for analysis. Ryder collected NHL game summaries into a database, though only 90% of games were available for the 2002–2003 season that was analyzed.

The game summary data were manually recorded by Real Time Scoring System (RTSS) trackers. For every shot, these trackers manually entered the player taking the shot, the approximate distance of the shot, and the type of shot (e.g., slap shot, wrist shot, etc.). Given the subjective human component in the RTSS tracking, Schuckers (2014) and others attempted to define consistent and systematic errors within the tracking data. For example, although shots are events that are relatively consistently recorded from rink to rink, St. Louis and Florida have shot rates that are significantly lower and higher, respectively, than other rinks in the league. Differences between rinks in terms of data tracking are exacerbated for other game events, such as giveaways, takeaways, hits, and blocked shots. These differences motivate data sites such as HockeyViz (**http://hockeyviz.com/**) to make rink adjustments for many of these metrics.

The original shot quality calculation by Ryder in 2004 did not make systematic rink adjustments beyond removing obvious errors in shot location. After surveying various shot types as well as the effect of rebounds and power plays, Ryder fitted goal probability curves by bucketing each shot type into 5-foot intervals. These goal probabilities were then summed to create the first expected goals against metric for hockey. Ryder also crucially identified that setting an expectation for goals against based on averaged win probabilities would allow us to evaluate goaltending ability while accounting for differences between goaltenders in terms of workload and team defense.

There have also been efforts to develop expected goals metrics without evaluating the quality of individual shots. Brian Macdonald (2012) created an expected goals framework using ordinary least squares and ridge regression models that were trained on several game events, namely the rates of goals, shots, hits, hits against, and face-offs. Ridge regression was used to address collinearity, which is a fundamental problem in almost any regression approach to hockey since subsets of players share significant ice time together as part of set lines.

With three seasons of training data, these models were able to more accurately predict future even strength (5 players versus 5 players) goals for an entire testing season's games compared with using any of those game events in isolation. Using this expected goals statistic as a response variable for an adjusted plus-minus model, Macdonald arrived at a rate statistic that summarized the offensive contribution of individual players in terms of even strength goals.

Dawson Sprigings and Asmae Toumi (Sprigings & Toumi 2015) created an expected goals regression model that also takes into account shot quality, as measured by shot type, angle, distance from goal, and the player's stronger side. They also made the interesting decision to incorporate shooting talent, the idea that some players have inherently higher chances of scoring on the same type/location of shot, by including a shot multiplier for each player. In their method, a player's

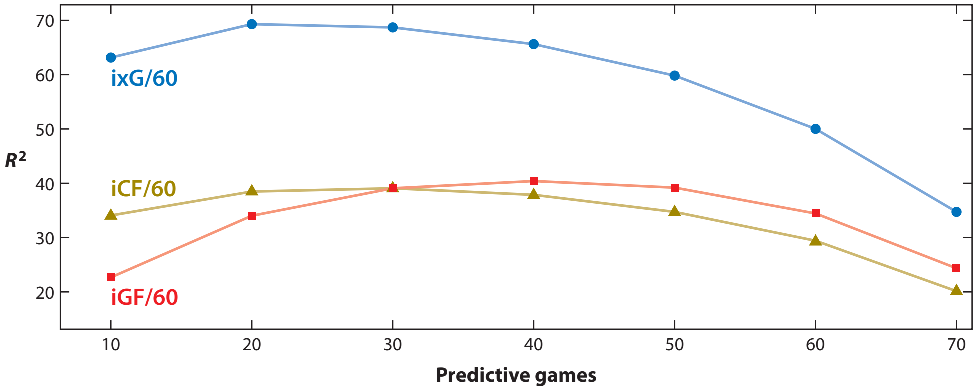


Figure 2

R^2 validation of individual expected goals. The x-axis refers to the number of games in the training set, while the remaining games from the 82-game season were used for testing. ixG/60, iCF/60, and iGF/60 correspond to individual expected goals for, Corsi for, and goals for per 60 respectively, with expected goals showing superior predictive ability in terms of R^2 for any sample size. Adapted with permission from A. Toumi.

shooting percentage is regressed to the mean shooting percentage by a factor that depends on the number of shots taken by a player within a season. By splitting team-seasons into randomly assigned training and testing sets of varying lengths, they were able to validate that with 20 or more games observed, their expected goal values are more predictive of future goals than Corsi or goals for alone (Sprigings & Toumi 2015). Similar results were obtained for players, in terms of predicting their on-ice goals and shots as well as their individual scoring, as shown in **Figure 2**.

A popular and publicly available expected goals model created by Manny Perry (2016a) uses similar shot quality predictor variables but does not incorporate shooting talent of individual players since individual shooting talent is difficult to estimate for players with good linemates. Each shot type is further subsetted into rebound and nonrebound shots, and then location, angle, etc. were used to predict goal probability using a logistic regression for each shot category.

Although there is no appreciable improvement over shot metrics in terms of predicting future on-ice goals for players, this approach does show improved predictions of individual scoring events and gives insight into what defines dangerous versus nondangerous shots.

The website Natural Stat Trick (**https://www.naturalstattrick.com/**), as well as the now-defunct website WAR On Ice, used data about shot quality to categorize shots into low-, medium-, and high-danger scoring chances. Although this binning of shots into discrete danger zones is a simplification of a more continuous process, it does allow for ease of interpretation, e.g., "team X had a lot of shots but team Y had more really good chances" (Thomas 2014b).

To summarize, current expected goals models emphasize the importance of measuring shot quality through angle, location, shot type, etc., but differ in the degree to which they claim predictive superiority over simpler shot metrics for season-level trends. While some sites, such as Natural Stat Trick, discretize shots into danger categories, there is a loss of information in this discretization compared with weighting shots continuously in these expected goals frameworks (Perry 2016b).

Overall, models that incorporate shot information have enjoyed considerable recent popularity, in part because the issue of shot quality feels instinctively important to analysts, coaches, players, and fans alike. For example, it is unlikely that coaches will ever tell their players to emphasize quantity of shots over quality of shots.

6. GOALTENDERS

One position in hockey that must be treated in a unique way is that of the goaltender, whose entire role is the prevention of goals against their team. It is worth noting that statistics about the goalie are missing from the traditional box score shown in **Figure 1**.

A criticism of most basic goaltending statistics is that they reflect the performance of the team as a whole rather than the goaltender specifically. The number of goals allowed within a game is a function of the goaltender's skill but also of the defense in front of him and their ability to prevent scoring chances, as well as the offense in front of them and their ability to maintain possession of the puck. The same can be said of unadjusted save percentage, although this metric does at least account for the number of shots faced by the goalie.

The most notable development in evaluating goaltenders has been the expected goals frameworks reviewed in the previous section. Being able to judge the expectation of any individual shot allows us to evaluate goaltender performance relative to the hypothetical average NHL goaltender. This leads us to the most popular current advanced statistics for goalies: adjusted save percentage and goals saved above average (Balloch 2015).

Adjusted save percentage represents the difference between a goalie's actual results and how a hypothetical league average goalie would perform in the same situation, which accounts for the strength of defenses and other contextual factors. Goals saved above average converts that adjusted save percentage into goals by accounting for the quantity of shots that particular goaltender has faced over any given time period. Many versions of these statistics have been calculated based on various expected goals frameworks; one of the earliest was created by Desjardins (2007) by binning shots based on location and calculating the empirical probabilities of resulting in a goal.

Subsequent advancements in these goaltender metrics have involved modifications to the underlying expected goal frameworks, such as breaking down adjusted save percentage into low-danger, medium-danger, and high-danger save percentages (Way 2017). One interesting finding is that most NHL-caliber goalies have relatively little trouble with low- and medium-danger shots, while the separation between goaltenders is seen in their performance on high-danger shots, which are defined by Corsica as shots with a conversion rate of 9% or greater (Perry 2016b).

7. POINT SHARES AND WINS ABOVE REPLACEMENT

All of the preceding hockey analytic metrics that we have reviewed are focused on estimating player performance in specific facets of the game. However, there is also a demand for more comprehensive performance measures that capture the entirety of a player's contributions to their team's chances of winning.

One of the most frequent criticisms of hockey analytics is the lack of appropriate emphasis on quality of defense because defense is more difficult to empirically measure and compare. There is also an anecdotal belief that goaltenders have the most impact on any given game, with the potential to steal a win for their team even if they have been outshot by wide margin.

For major league baseball, WAR has become the most popular comprehensive measure of player performance by accounting for player contributions to hitting, fielding, pitching, and base running (Baumer et al. 2015). It has been more difficult to extend the ideas of WAR to other sports such as hockey because each aspect of game play is more difficult to isolate and measure quantitatively. There are various aggregate measures of performance for hockey, but no single one has achieved the same dominance as WAR in baseball.

Thomas (2014a) of WAR On Ice provided extensive detail on the considerations involved in creating a WAR metric for hockey. After establishing that the ideal metric should be predictive and linearly decomposable, he argued that the optimal metric should describe a player's impact on any event that is predictive of goals being scored. Unfortunately, his methodology and WAR values are no longer available to the public.

Ryder (2004b) had one of the earliest methods for assigning individual credit to players for team outcomes, heavily inspired by Bill James's win shares in baseball (James & Henzler 2002). He starts by calculating marginal goals scored per team, which is simply goal differential + league average goals scored, and also introduces the concept of a marginal, or near replacement level, hockey player. These marginal goals are decomposed into marginal goals prevented and created, allocated to different strength states, and, finally, attributed to individual players. For example, offensively, 50% of each goal is given to the scorer, with the rest of the credit given to the playmakers who assisted. While the component parts of hockey performance are defined in a principled way, Ryder's approach does run into the issues of using actual goals rather than expected, and relying on multipliers that have generally not been validated.

Kubatko (2011) uses a similar methodology with different multipliers to calculate a metric called Point Shares and notes that summed Point Shares by team differ from their actual season outcome by 5 standings points on average. Goals versus threshold, developed by Awad (2009), also falls within this general methodological category, though it also accounts for player shootout value, which is not commonly done.

Dom Luszczyszyn's Game Score statistic is a popular and publicly available comprehensive metric (Luszczyszyn, 2016) that is available on hockey statistics sites (**http://corsica.hockey**). Game Score was originally developed by Bill James (1988), and brought to basketball by John Hollinger (2010). Luszczyszyn's (2016) Game Score attempts to assess game-level productivity by weighting box score statistics by their frequency to goals. For skaters, Game Score ends up being a linear combination of goals, primary assists, secondary assists, shots on goal, blocked shots, penalty differential, face-offs, 5-on-5 shot differential, and 5-on-5 goal differential. Game Score values are fairly consistent from year to year and easy to construct and interpret.

The popularity of these metrics certainly depends on the quality and transparency of the modeling effort, but it also depends heavily on availability. Game Score, for example, is readily found on corsica.hockey, and so can easily be mentioned in conjunction with shot-based metrics in player evaluations. Similarly, Point Shares are updated daily on **https://www.Hockey-Reference.com**, as are the latest WAR metric and ensemble player ratings on corsica.hockey.

8. PLAYER AGING AND CONTRACT VALUATION

With any measure of player value, such as Point Shares or Game Score or even plus-minus, there is an obvious interest (and substantial recent discussion) in how those values change as a function of age, and evaluation of the extent to which we can forecast player value into the future. One recent finding is that players tend to peak in their early 20s, which differs from both the conventional wisdom and the NHL salary structure that has players making below-market value until they are able to hit free agency in their late 20s (Solberg & Solberg 2017).

Future forecasts of performance are absolutely crucial for contract valuation, which is a major component of hockey research for teams as well as the broader public community. Simple early measures like cost per goal and cost per point have mostly given way to more rigorous statistical projections. For example, Cane (2017) uses a random forest model to predict a range of possible values for the average annual value of a contract, given what general managers have offered comparable players in the past.

The discussion of forecasting into the future and contract valuation also moves the discussion of hockey analytics into the realm of discovering and exploiting market inefficiencies within the NHL. There is perhaps no better setting in which to search for these inefficiencies than the NHL entry draft.

9. SPECIAL FOCUS ON DRAFT STRATEGY AND EVALUATION

The NHL draft currently comprises seven rounds, and each team is initially allotted one pick per round. The importance of drafting talent versus acquiring it through free agency cannot be overstated. The NHL has a hard salary cap of $75,000,000 for the 2017–2018 season, and the best players can command contracts of upwards of $10,000,000 per year. In order to fill out a roster of 23 or so players, it is necessary to find value elsewhere, particularly with younger players who are cost controlled. Entry-level contracts cover the first three seasons of every player's career at a maximum of $925,000 per year (plus performance bonuses) in addition to four years of restricted free agency.

At a team level, there is no evidence that draft success is consistent across years. Teams do, however, seem to outperform publicly available scouting data at selecting players (Schuckers & Argeris 2015). To improve upon the NHL's draft ordering of prospects, Schuckers & Argeris (2015) use a generalized additive model trained on demographic information as well as junior league performance metrics to predict NHL games played and time on ice in the first seven years of a player's career, which is the time period for which their draft team has their exclusive rights.

Defining drafting success is complicated by whether to assume a team-specific or league-wide perspective. Schuckers & Argeris (2015) assume that the outcome of interest is overall deviation from optimal draft ordering, and success would entail the league as a whole drafting players in descending order of future NHL value.

As Nandakumar (2017a) shows, moving to a team-specific perspective requires a shift in evaluation in order to integrate game theoretic principles. The best draft decision becomes picking the best possible player with the latest possible pick, in order to maximize the return on investment of each pick and the overall prospect value found throughout the draft.

Under Nandakumar's (2017a) framework, NHL teams were evaluated based on their best possible outcomes, given the number and value of their picks, within each draft from 2000 to 2009. Averaged rankings of teams in terms of draft efficiency (total value drafted divided by possible value in terms of Hockey Reference's Point Shares) did not differ significantly from what would be expected if team rankings varied randomly from year to year. **Figure 3** shows this distribution of averaged ranks, with perhaps the most notable pattern being a slight separation between the top 20 and bottom 10 teams. In general, it is difficult to discern a signal in terms of team proficiency in drafting because of the randomness inherent in all aspects of the draft.

One of the most well explored topics in draft research across sports is defining the relative value of different draft slots. How much less is a second overall pick worth than a first, and does this relationship hold through each round? These questions are particularly enticing as historical data are readily available, and there are major implications for how teams construct trades around draft picks.

In the National Football League (NFL) teams use a well-known chart to evaluate potential trades, developed by the Dallas Cowboys in the early 1990s (Smith 2017). The chart shows a roughly exponential decay in pick value. Hurley et al. (2012) constructed a negative binomial distribution model with three classes of NFL players, and they find that the probability of someone belonging to the class with the most NFL games played decreases as the player is drafted later. However, Hurley et al.'s (2012) construction of "The Chart" of pick valuation substantially differs

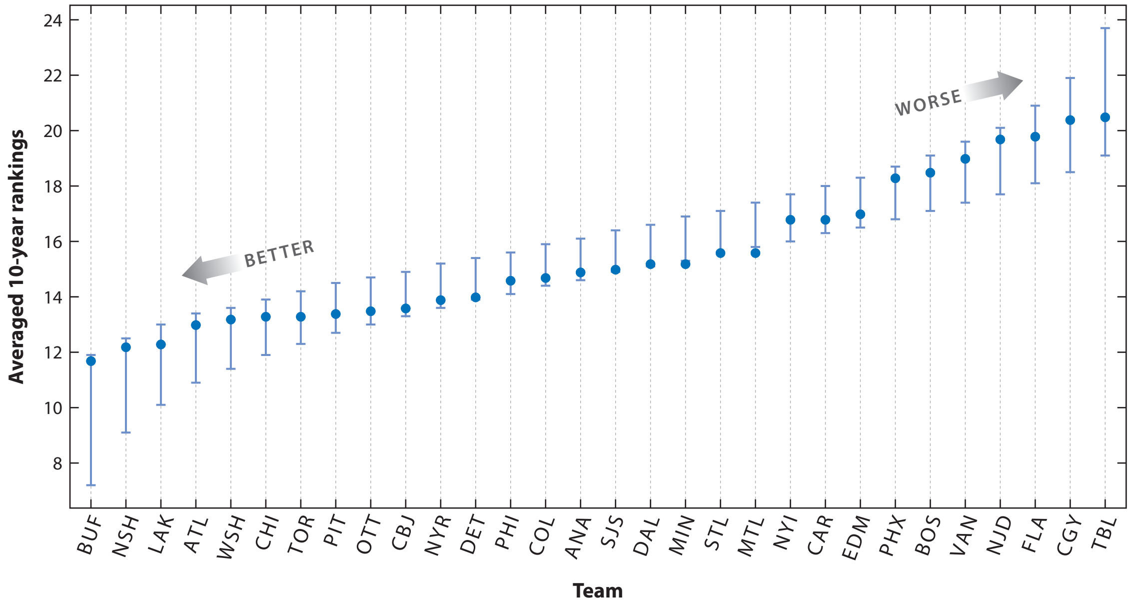


Figure 3

Distribution of draft efficiency ranking for each National Hockey League team across 10 years of drafting. Adapted with permission from N. Nandakumar.

from the original chart of Jimmy Johnson's Cowboys team, suggesting that conventional wisdom is not accurately capturing the way in which value declines with declining draft slot.

In terms of analysis of the NHL draft, Schuckers (2011) and Tulsky (2013a) construct draft pick value charts with similarly decaying trends. **Figure 4** shows Schuckers's most recently updated value chart (Schuckers 2016). The common framework for evaluating draft picks across all sports is averaging the values of players picked at each pick over a number of years, and then smoothing

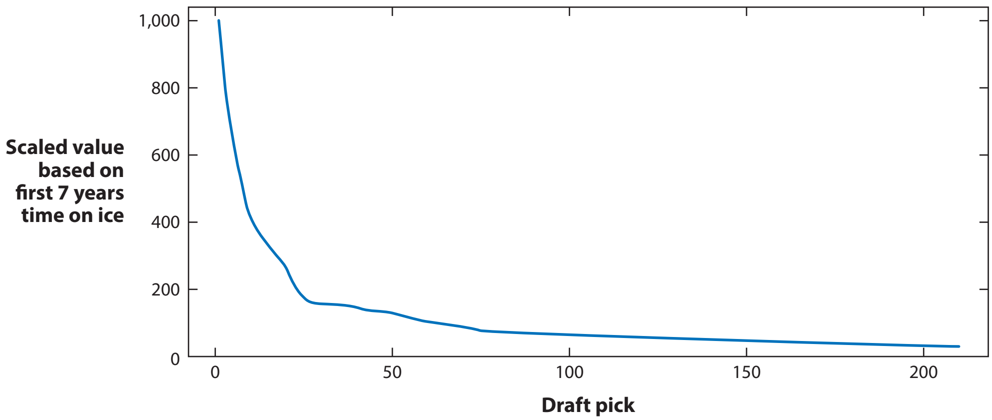


Figure 4

Chart of the average Point Share value of players selected at each draft slot of the NHL draft (adapted with permission from Schuckers 2016).

the trend. However, as we have discussed above, assigning an appropriate value to each player is a complex and difficult task.

In particular, in order to improve draft analyses for hockey, it will be necessary to move beyond player value measures such as NHL games played and instead use some of the more comprehensive metrics that we outline in Section 7. Using games played as the measure of career success in any type of draft analysis may prioritize the selection of safe players who can fill out a roster but do not necessarily have significant NHL value on a per-game basis. A quick example can be seen in the comparison of the Washington Capitals' two first-round picks in 2012, Tom Wilson and Filip Forsberg. Wilson has 60 more regular season games played as of the 2017–2018 season, but by every other measure of player value from career points to average WAR, Forsberg is an overall more impactful player. To cite a value metric mentioned previously, Forsberg's career Game Score per 60 minutes is 3.13, more than double Wilson's mark of 1.32, per corsica.hockey (**https://www.corsicahockey.com/nhl/players/nhl-player-stats**).

Moving beyond naive metrics for career performance, such as number of games played, is easier said than done. One of the obvious limitations of more advanced measures, such as the comprehensive metrics from Section 7, is that these measures are only consistently available from the 2007–2008 season onward (Vollman 2016). Since most drafted players take at least a few years to even begin their NHL careers (Nandakumar 2017c), we currently have a limited number of recent drafts for which we can use more comprehensive measures of career performance.

Figure 5 shows that roughly one third of drafted prospects eventually play 80 NHL games, which could be considered a reasonable sample from which to draw conclusions of player value. However, it takes at least 574 NHL games, or 7 seasons, after draft day for the survival curve to flatten out such that we are reasonably sure that a player who has not yet hit the 80-game benchmark will not do so in the future. As of the conclusion of the 2017–18 NHL regular season, there are only 5 draft years (2007–2011) that have advanced player performance metrics available as well as at least 7 full seasons of data. Fortunately, the group of draft years that can be analyzed

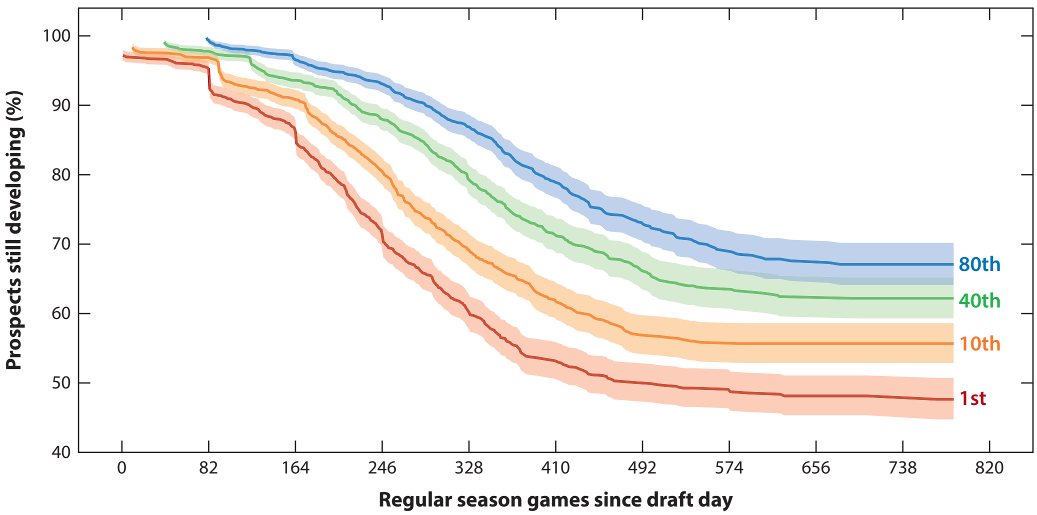


Figure 5

Curves of time until *n*th career NHL game. For each length of time (in terms of National Hockey League games) since draft day, we give the percentage of drafted prospects who have played their 1st, 10th, 40th, and 80th game by that point in time.

with relatively sophisticated value metrics will continue to widen in the future, as more seasons of advanced metrics become available.

Schuckers (2016) mentions another weakness of existing draft analyses: Current draft slot value charts do not imply that value of draft slots must be monotonically decreasing, which should theoretically be the case. A draft pick slot should always be at least as valuable as the one directly following it, since the number of players available to the team is always larger with an earlier pick. Perhaps future studies of draft pick value could explore how substantially the value across the prospect choice set decreases after each draft pick, which could lead to the construction of some sort of draft pick opportunity chart.

At a player level, there is evidence to suggest that teams are undervaluing prospects who are shorter, non-North American, and have later birthdays relative to their cohort. Deaner et al. (2013) show evidence of a bias against hockey players born in later months because of how age cutoffs work in lower levels of the sport. Fyffe (2009) has done extensive research on the relationship between factors like age, junior league point production, and more on various NHL draft outcomes.

NHL equivalencies were among the first publicly available measures used to evaluate individual prospects in a predictive sense (Desjardins 2005). These equivalencies are multipliers that use a prospect's point production to project how many points they would score as an NHL player. This approach is attractive in its simplicity: Anyone can make relatively educated guesses about a prospect's potential with knowledge of the appropriate multipliers. Of course, there are many additional variables about prospects that should be taken into account in order to more precisely project draft outcomes.

Perhaps the best known public example of an overarching statistical model for evaluating draft prospects is the Prospect Cohort Success (PCS) model developed by Weissbock (2015) to address some of the common deficiencies in NHL prospect analysis. This is also one of the best examples of the inherent risk in public hockey analytics research. The Florida Panthers hired the creators of PCS as scouting consultants, and thus the tool was lost to the public, although some documentation remains (Fialkov 2016). As analysts build on descriptive historical research by using more sophisticated predictive techniques for draft analysis, the probability of teams purchasing these tools and removing them from the public sphere grows ever higher. While the field has certainly not reached the heights seen in baseball after the popularity of *Moneyball*, most NHL teams have at least entertained the value of analytical tools in terms of assisting hockey operations decision making.

We do not know the exact methodology of PCS, but we do know that the authors made a critical decision to explicitly model the chance that a drafted prospect does not actually have any substantive impact in the NHL. Instead of a singular unified metric, Weissbock (2015) looked at both the probability of making it to the NHL (defined as playing 200+ NHL games) as well as the estimated value (defined as points per NHL game) of players who do make it to the NHL.

This decision is worthy of further consideration, as it seems to address some of the concerns expressed earlier with use of games played alone as a response variable. Potential to become an NHL player and projected NHL value are correlated but not equivalent. As Nandakumar (2017c) discusses, making it to the NHL requires significant buy-in from the team in question, as they decide which prospects are given the initial opportunity to play and develop. More than half of all drafted players never play a single NHL game. Once a player manages to break into a roster, they are free to perform to the fullest of their ability, although projection of value for full time players has all the difficulties previously discussed.

A similar confounding issue is seen in projecting draft outcomes versus projecting NHL outcomes. A natural framework for diagnosing under- and overvalued traits is seen in "From college to the pros: predicting the NBA amateur draft" by Berri et al. (2011). The authors first construct a model to predict the draft position of a prospect based on factors such as height, college performance, etc. They then go on to construct an additional model to predict the National Basketball Association (NBA) performance of a draft pick. The disparities in the weighting of predictors across the two models shed light as to which traits are under- and overvalued. Mulholland & Jensen (2014) compare models for predicting draft order versus predicting career performance for tight ends in the NFL. They also find disparities in the choice of predictors for draft order versus career performance, which suggests some traits are overvalued by the draft.

While these comparisons are insightful, it is undeniable that the draft outcomes themselves are an excellent predictor of NHL outcomes in large part because they signal a team's level of investment in a prospect. There is no shortage of evidence to suggest that teams are particularly hesitant to give up on their high-value picks and may afford them more opportunities to succeed.

There is also the question of the boom-or-bust prospect seen across sports, i.e., who may have an unusually low probability of making it but an unusually high projected value. Nandakumar (2017b) shows that based on the salary cap structure of the NHL, where successful prospects are extra valuable in terms of having a small cap hit for their entry-level contract, boom-or-bust prospects become relatively more attractive. The preferred outcome variable becomes a prospect producing value above and beyond what can be found by other means, including trades and free agency. Furthermore, in a league where the salary cap provides a significant constraint to most teams, tying this prospect value to contract dollars is a meaningful endeavor, which is also shown by Massey & Thaler (2013) for the NFL.

Accounting for the inherent uncertainty in prospect success, as well as the role of teams in shaping outcomes for drafted players, can improve draft research done at an individual level, while explicitly defining and controlling for incentives can improve draft research done at a team level.

10. FUTURE DIRECTIONS AND RECOMMENDATIONS

One institutional change that could completely revolutionize the landscape of hockey analytics would be an increased effort to integrate player tracking technology at all levels of the sport. One underlying problem of assigning credit to hockey players is the dearth of information regarding players who are not directly involved with scoring or stopping any particular goal.

In the absence of any initiative by the NHL to implement a league-wide tracking system, communities of trackers have sought to collect the data themselves. Ryan Stimson's Passing Project is perhaps the best known of these efforts, requiring considerable individual commitment to generate the data at a much slower and more mistake-prone rate than any tracking system (Goldman 2016).

These efforts serve to provide additional context in areas that are not given by traditional metrics available through the NHL (goals, assists, shots, time on ice, etc.). One example is the idea of shot assists, giving credit to playmakers who set up their teammates even when those teammates are not able to score. Another direction is the success rate and types of passes used by players (Stimson & Cane 2017). Yet another is the concept of successful zone entries and exits. Players who are able to exit their defensive zone and enter the offensive zone with the puck under control can more consistently drive play and create scoring opportunities for their teams (Luszczyszyn 2015).

Despite the difficulties inherent in data collection, tracking data have already generated valuable insights for the hockey analytics community. They have helped to close the gap between hockey researchers, who employ statistics and macro-level analysis, and players and coaches, who are concerned with the efficacy of hockey systems and more granular strategies.

While tracking data are not a cure-all solution to player evaluation, other sports have made considerable advances due to the availability of tracking data. The most salient example is the NBA, which has released tracking data to media outlets and the public (Barker 2016). The NFL is also following along with a recent commitment to disseminating league-wide tracking data to all teams.

A specific area of on-ice performance worthy of further public study is power play and short-handed situations. There is a tendency to reference commonly used statistics such as Corsi, expected goals, etc., solely in terms of 5v5 play, and for good reason. Special teams opportunities (when one team has the advantage of more players on the ice) vary substantially from game to game, and very few players per team are consistently deployed in these situations. These situations can often be the difference between winning and losing, and being able to define the extent to which player and team prowess matter in special teams situations is a worthy endeavor.

A further recommendation in this direction would be for hockey analysts to use more intuitive naming conventions for new and modified hockey statistics. It is something of a running joke that NHL players and coaches extol the virtues of getting pucks to the net and, in the same breath, diminish the importance of Corsi. The reality, of course, is that Corsi directly measures "getting pucks to the net," and could just as easily be referenced as "shots" or "shot attempts." The entire sports analytics movement suffers in terms of branding when its aims are perceived as being in opposition to teams and coaches rather than complementary to conventional wisdom.

There has been a movement within sports analytics toward prioritizing proprietary data and methodology at the expense of reproducible, well-validated methods. For the community, it is remarkably inefficient to have to continually rediscover existing results and techniques because they are not publicly known. Moving forward, there should be an increased dialogue about data and results sharing in such a way that still maintains team incentives for gaining a competitive advantage.

Despite the general trend of resources going dark after their creators are picked up by NHL teams, people interested in public hockey analytics do still have a wealth of resources at their disposal. For example, **https://www.NHL.com** lists standard box score statistics such as goals, assists, and time on ice, with some shot reports and other breakdowns available upon further navigation of the site. However, player-level and game-level shot metrics are more readily available at Hockey Reference, corsica.hockey, and Natural Stat Trick. Hockey Reference features basic shot metrics such as Corsi and Fenwick, in addition to their Point Shares value metric. Natural Stat Trick publishes scoring chances to address the question of shot quality differences, while corsica.hockey has its own expected goals model and proprietary player rating system. Another site, **https://www.moneypuck.com**, publishes real-time in-game results of its expected goals model and win probabilities. Additionally, HockeyViz provides visual summaries of player and team performance, including maps that show shot location tendencies of players and teams. Finally, for those wishing to read and iterate off of existing literature, the database MetaHockey has conveniently aggregated the majority of public hockey analytics resources.

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

LITERATURE CITED

Awad T. 2009. Understanding GVT, part 3. *Hockey Prospectus*, Aug. 4

Balloch G. 2015. An introduction to adjusted save percentage. *InGoal Magazine*, Sep. 7. **http://ingoalmag.com/analysis/an-introduction-to-adjusted-save-percentage/**

Barker M. 2016. The NBA is wisely opening up its SportVU data to fans and media outlets. *Sport-Techie*, Jan. 20. **https://www.sporttechie.com/the-nba-is-wisely-opening-up-its-sportvu-data-to-fans-and-media-outlets/**

Baumer BS, Jensen ST, Matthews GJ. 2015. OpenWAR: an open source system for evaluating overall player performance in major league baseball. *J. Quant. Anal. Sports* 11:69–84

Berri DJ, Brook SL, Fenn AJ. 2011. From college to the pros: predicting the NBA amateur draft. *J. Product. Anal.* 35(1):25–35

Cane M. 2014. Breaking down Corsi: looking into shot blocks and misses. *Puck++ Blog*, Feb. 11. **https://puckplusplus.com/2014/02/11/breaking-down-corsi-looking-into-shot-blocks-and-misses/**

Cane M. 2017. *Predicting NHL player contracts*. Presentation at the RIT Hockey Analytics Conference, Rochester, NY, Oct. 21

Deaner RO, Lowen A, Cobley S. 2013. Born at the wrong time: selection bias in the NHL draft. *PLOS ONE* 8(2):e57753

Desjardins G. 2005. Projecting junior hockey players and translating performance to the NHL. *Behind the Net*, May 26. **http://www.behindthenet.ca/projecting_to_nhl.php**

Desjardins G. 2007. 2007–08 5v5 Goaltender performance. *Behind the Net Blog*, Dec. 23. **http://www.behindthenet.ca/blog/2007/12/2007-08-5v5-goaltender-performance.html**

Desjardins G. 2009a. Advanced metric frequently asked question #1: A quick spin through +/- statistics at Behind the Net. *Arctic Ice Hockey Blog*, Oct. 8. **https://www.arcticicehockey.com/2009/10/8/1076067/frequently-asked-question-1-a**

Desjardins G. 2009b. Advanced metric frequently asked question #2: What does quality of competition tell us about NHL players? *Arctic Ice Hockey Blog*, Oct. 8. **https://www.arcticicehockey.com/2009/10/8/1076586/frequently-asked-question-2**

Desjardins G. 2009c. Advanced metric frequently asked questions #4: Zone starts (offensive-to-defensive face-off ratio). *Arctic Ice Hockey Blog*, Oct. 8. **https://www.arcticicehockey.com/2009/10/8/1077082/frequently-asked-questions-4-zone**

Desjardins G. 2009d. Advanced metric frequently asked questions #5: Penalty +/-. *Arctic Ice Hockey Blog*, Oct. 8. **https://www.arcticicehockey.com/2009/10/8/1077182/frequently-asked-questions-5**

Desjardins G. 2009e. Advanced metric frequently asked question #9: Quality of teammates. *Arctic Ice Hockey Blog*, Oct. 9. **https://www.arcticicehockey.com/2009/10/9/1078388/frequently-asked-question-9**

Desjardins G. 2009f. Advanced metric frequently asked questions about statistical analysis in the NHL. *Arctic Ice Hockey Blog*, Oct. 9. **https://www.arcticicehockey.com/2009/10/9/1078607/frequently-asked-questions-about**

Fialkov H. 2016. Panthers enter analytical era of drafting. *Sun-Sentinel*, Jun. 23. **http://www.sun-sentinel.com/sports/florida-panthers/fl-florida-panthers-0624-20160623-story.html**

Fyffe I. 2009. Up and coming. *Hockey Prospectus*, May 19. **http://hockey.dv2013.baseballprospectus.com/article.php?articleid=135**

Goldman S. 2016. Phasing into analytics: a primer on NHL analytics and how experts use them (part 1). *Sport-Techie*, Oct. 12. **https://www.sporttechie.com/phasing-into-analytics-a-primer-on-nhl-analytics-and-how-to-use-them-part-1/**

Gramacy RB, Jensen ST, Taddy M. 2013. Estimating player contribution in hockey with regularized logistic regression. *J. Quant. Anal. Sports* 9:97–111

Hollinger J. 2010. Best finals performances: the method. *ESPN*, Jun. 11. **http://www.espn.com/nba/playoffs/2010/columns/story?columnist=hollinger_john&page=BestFinalsGame-Individual-intro**

Hurley WJ, Pavlov A, Andrews W. 2012. A valuation model for NFL and NHL draft positions. *J. Oper. Res. Soc.* 63(7):890–98

James B. 1988. *The Bill James Baseball Abstract*. New York: Ballantine

James B, Henzler J. 2002. *Win Shares*. Chicago: STATS Inc.

Kubatko J. 2011. Calculating Point Shares. *Hockey Reference*, Mar. 12. **http://www.hockey-reference.com/about/point_shares.html**

Lewis M. 2004. *Moneyball: The Art of Winning an Unfair Game*. New York: Norton

Luszczyszyn D. 2015. Two years of research, 1230 games watched: why neutral zone play is much more important than you think. *The Hockey News*, Sep. 29. **http://www.thehockeynews.com/news/article/two-years-of-research-1230-games-watched-why-neutral-zone-play-is-much-more-important-than-you-think**

Luszczyszyn D. 2016. Measuring single game productivity: an introduction to Game Score. *Hockey Graphs*, Jul. 13. **https://hockey-graphs.com/2016/07/13/measuring-single-game-productivity-an-introduction-to-game-score/**

Macdonald B. 2012. *An expected goals model for evaluating NHL teams and players*. Paper presented at the MIT Sloan Sports Analytics Conference, Boston, MA, Mar. 2–3. **http://www.sloansportsconference.com/wp-content/uploads/2012/02/NHL-Expected-Goals-Brian-Macdonald.pdf**

Massey C, Thaler RH. 2013. The loser's curse: decision making and market efficiency in the National Football League draft. *Manag. Sci.* 59(7):1479–95

McCurdy MB. 2014. Adjusted possession measures. *Hockey Graphs*, Nov. 13. **https://hockey-graphs.com/2014/11/13/adjusted-possession-measures/**

McKenzie B. 2014. The real story of how Corsi got its name. *TSN*, Oct. 6. **https://www.tsn.ca/mckenzie-the-real-story-of-how-corsi-got-its-name-1.100011**

Mulholland J, Jensen ST. 2014. Predicting the draft and career success of tight ends in the National Football League. *J. Quant. Anal. Sports* 10:381–96

Nandakumar N. 2017a. *What does it mean to draft perfectly in the NHL?* Presented at the Vancouver Hockey Analytics Conference, Vancouver, Canada, Mar. 11

Nandakumar N. 2017b. Exploiting variance in the NHL draft. *Hockey Graphs*, Mar. 28. **https://hockey-graphs.com/2017/03/28/exploiting-variance-in-the-nhl-draft/**

Nandakumar N. 2017c. *Survival analysis of NHL prospect timelines*. Presentation at the RIT Hockey Analytics Conference, Rochester, NY, Oct. 21

Perry E. 2016a. Shot quality and expected goals: part I. *Corsica*, Mar. 3. **http://www.corsica.hockey/blog/2016/03/03/shot-quality-and-expected-goals-part-i/**

Perry E. 2016b. Shot quality and expected goals: part 1.5. *Corsica*, Aug. 13. **http://www.corsica.hockey/blog/2016/08/13/shot-quality-and-expected-goals-part-1-5/**

Ryder A. 2004a. *Shot quality*. Work. Pap., Hockey Analytics. **http://hockeyanalytics.com/Research_files/Shot_Quality.pdf**

Ryder A. 2004b. *Player contribution*. Work. Pap., Hockey Analytics. **http://www.hockeyanalytics.com/Research_files/Player_Contribution_System.pdf**

Salfino M. 2017. Blake Bortles is the Tom Brady of garbage time. *FiveThirtyEight*, Aug. 16. **https://fivethirtyeight.com/features/blake-bortles-is-the-tom-brady-of-nfl-garbage-time/**

Schuckers ME. 2011. *What's an NHL draft pick worth? A value pick chart for the National Hockey League*. Work. Pap., St. Lawrence Univ. and Stat. Sports Consult., Canton, NY

Schuckers ME. 2014. Dealing with the NHL's RTSS recording of events. *Statistical Sports Consulting*, Nov. 24. **http://statsportsconsulting.com/2014/11/24/1357/**

Schuckers ME. 2016. Draft pick value chart. *Statistical Sports Consulting*, May 26. **http://statsportsconsulting.com/2016/05/26/1658/**

Schuckers ME, Argeris S. 2015. You can beat the "market": estimating the return on investment for NHL team scouting. *J. Sports Analytics* 1:111–19

Schuckers ME, Lock DF, Wells C, Knickerbocker CJ, Lock RH. 2011. *National Hockey League skater ratings based upon all on-ice events: an adjusted minus/plus probability (AMPP) approach*. Tech. Rep., St. Lawrence Univ., Canton, NY. **http://myslu.stlawu.edu/~msch/sports/LockSchuckersProbPlusMinus113010.pdf**

Smith MD. 2017. NFL teams have revised the draft trade chart. *NBC Sports/ProFootballTalk Blog*, Apr. 30. **http://profootballtalk.nbcsports.com/2017/04/30/nfl-teams-have-revised-the-draft-trade-chart/**

Solberg L, Solberg J. 2017. A new look at aging curves for NHL skaters (part 1). *Hockey Graphs*, Mar. 23. **https://hockey-graphs.com/2017/03/23/a-new-look-at-aging-curves-for-nhl-skaters-part-1/**

Sprigings D, Toumi A. 2015. Expected goals are a better predictor of future scoring than Corsi, goals. *Hockey Graphs*, Oct. 1. **https://hockey-graphs.com/2015/10/01/expected-goals-are-a-better-predictor-of-future-scoring-than-corsi-goals/**

Stimson R, Cane M. 2017. *Evaluating defensive ability in hockey using passing data*. Paper presented at the MIT Sloan Sports Analytics Conference, Boston, MA, Mar. 3–4. **http://www.sloansportsconference.com/wp-content/uploads/2017/02/1614.pdf**

Thomas AC. 2014a. The road to WAR (for hockey), part 1: the single-number dream. *WAR On Ice: The Blog*, Oct. 5. **http://blog.war-on-ice.com/index.html%3Fp=37.html**

Thomas AC. 2014b. NEW: Defining scoring chances. *WAR On Ice: The Blog*, Dec. 31. **http://blog.war-on-ice.com/new-defining-scoring-chances/index.html**

Thomas AC, Ventura SL, Jensen ST, Ma S. 2013. Competing process hazard function models for player ratings in ice hockey. *Ann. Appl. Stat.* 7:1497–524

Tulsky E. 2011. Simplify scoring: drop the pointless secondary assist. *Broad Street Hockey*, Mar. 15. **https://www.broadstreethockey.com/2011/3/15/2046512/simplify-scoring-drop-secondary-assists**

Tulsky E. 2012. Adjusting for score effects to improve our predictions. *Broad Street Hockey*, Jan. 23. **https://www.broadstreethockey.com/2012/1/23/2722089/score-adjusted-fenwick**

Tulsky E. 2013a. NHL draft: What does it cost to trade up? *Broad Street Hockey*, Apr. 25. **https://www.broadstreethockey.com/2013/4/25/4262594/nhl-draft-pick-value-trading-up**

Tulsky E. 2013b. Shooting percentage regression. *Broad Street Hockey*, Jun. 13. **https://www.broadstreethockey.com/2013/6/13/4427522/shooting-percentage-regression**

Vollman R. 2016. Complete list of hockey analytics data resources. *Hockey Abstract*, Sep. 15. **http://www.hockeyabstract.com/thoughts/completelistofhockeyanalyticsdataresources**

Way D. 2017. A beginner's guide to goaltender advanced stats. *Blueshirts Breakaway*, Nov. 16. **http://www.blueshirtsbreakaway.com/2017/11/16/a-beginners-guide-to-goaltender-advanced-stats/**

Weissbock J. 2015. Draft analytics: unveiling the Prospect Cohort Success model. *Canucks Army*, May 26. **https://canucksarmy.com/2015/05/26/draft-analytics-unveiling-the-prospect-cohort-success-model/**