

# Total Hockey Rating (THoR): A comprehensive statistical rating of National Hockey League forwards and defensemen based upon all on-ice events

© 2013 Michael Schuckers<sup>1,2</sup>, and James Curro<sup>1,3</sup>

<sup>1</sup>St. Lawrence University, <sup>2</sup>Statistical Sports Consulting LLC and <sup>3</sup>Iowa State University  
Canton, NY, USA, 13617,  
Email: [schuckers@stlawu.edu](mailto:schuckers@stlawu.edu)

## Abstract

Hockey is a fluid sport with players frequently coming on and off the ice without the stoppage of play. It is also a relatively low scoring sport compared to other sports such as basketball. Both of these features make evaluation of players difficult. Recently, there have been some attempts to get at the value of National Hockey League (NHL) players including Macdonald [1], Ferrari [2], and Awad [3]. Here we present a new comprehensive rating that accounts for other players on the ice will a give player as well as the impact of where a shift starts, often called zone starts [4], and of *every* non-shooting events such as turnovers and hits that occur when a player is on the ice. The impact of each play is determined by the probability that it leads to a goal for a player's team (or their opponent) in the subsequent 20 seconds. The primary outcome of this work is a reliable methodology that can quantify the impact of players in creating and preventing goals for both forwards and defenseman. We present results based on all events from the 2010-11 and 2011-12 NHL regular seasons.

## 1 Introduction

In this paper we present a novel comprehensive reliable methodology for the rating of National Hockey League (NHL) forwards (centers and wings) and defensemen. Our approach considers *every* event recorded by the NHL and assigns value to those events based upon the probability that they will lead to a goal. To evaluate players we determine which players were on the ice for which events and assess the impact of each player adjusting for their teammates and their opponents on the ice with them. Recent work has shown that where a player starts their shift (a shift in hockey is the continuous period when a player is on the ice) has an impact upon the events for which the player will be on the ice. This effect is known as Zone Starts and we explicitly model this effect as part of our ratings. Further, we include a home-ice effect. The result of all of this is the change in probability of a goal per event for each player. To facilitate comparisons we convert this number into wins above average for a season. Since this rating takes into account all of the events that occur when a player is on the ice, we refer to the ratings at the Total Hockey Ratings (THoR). In the rest of this paper we discuss the data involved in our analysis, our approach to analyzing these data and the results of our analysis.

## 2 Data

For this analysis we use data from the NHL's Real Time Scoring System (RTSS). That system records events that occur in every NHL game as well as which individuals were on the ice for those events. Specifically we use the Play by Play (PBP) files from every game to obtain the on-ice action events. These on-ice action events are: a faceoff (FAC), a hit (HIT), a giveaway (GIVE), a takeaway (TAKE), a blocked shot (BLOCK), a missed shot (MISS), a shot on goal (SHOT), a goal (GOAL), or a penalty (PENL). Other events such as stoppages, the beginning and ending of periods are not included for the evaluation of player performance. For each event we know where on the ice the event occurred. In the case of a HIT, FAC, TAKE, GIVE, MISS, BLOCK or PENL we know the zone on the ice, either Offensive, Defensive or Neutral, where the event occurred. For SHOT and GOAL events we also know the x and y coordinates from where the shot was taken.

Previous analyses of the NHL's RTSS data Ryder [4], Desjardins [5] and Fischer [6] have found biases in these data based upon the rink in which the data was collected. To account for these effects, we have made some adjustments to the event data. First recognizing the home rink bias in takeaways and giveaways, we lumped these two events together as turnovers (TURN). There is a negative connotation to a giveaway and a positive connotation to a takeaway which is likely responsible for the bias. By combining these two events we aim to negate the bias and to

2013 Research Paper Competition  
Presented by:

simply recognize the change of possession from one team to another. The second adjustment that we made was to the x- and y- coordinates for shots and goals. There is a bias in these shot location values at some rinks, most notably Madison Square Garden (MSG), home of the New York Rangers. In the case of MSG, shots taken there have a significantly different distribution than shots taken at other rinks. Previously, we utilized other adjustments as part of our previous work on Defense Independent Goalie Ratings, Schuckers [7]. Our adjustments here which are detailed further in Appendix I subtract the difference between the distribution of shots taken by all away teams at a rink R from the distribution of shots taken by all away teams at all rinks. We then adjust the distribution at rink R by this difference at the distribution level, i.e. the cumulative probability function level.

In addition to the event data, we have the name of all of the players who are on the ice for each event. We record this information as well as which players were on the ice for the home team for the event and which players were on the ice for the away team for the event. Note that for every event on the ice we will distribute value for that event to all of the players on the ice. This is done to account for the effect of an individual, for example, a Sidney Crosby or a Shea Weber, who may impact events but is not directly involved in a given event such as a shot. We also record the location where a particular shift starts. We will denote the variable ZS, for zone start, as a 1 if a shift starts in the home offensive zone, a zero if the shift starts in the neutral zone and a -1 if the shift starts in the home defensive zone. As several analysts including Charron [8] and Calloway [9] have noted, where an individual starts their shifts can inflate or deflate their offensive numbers. Our approach here is different than the zone start percentage that is typically used by hockey analysts which calculates the ratio of shift starts in the offensive zone to shift starts outside the neutral zone. The drawback to the latter statistic is that it treats a player with 40 defensive zone starts, 200 neutral zone starts and 60 offensive zone starts, a zone start percentage of 60%, the same as a player with 100 defensive zone starts, 50 neutral zone starts and 150 offensive zone starts. Our approach focuses on the number of additional starts in the offensive zone.

To assess the impact of each event we looked at the probability that it led to a goal. This was done primarily because of the low scoring rates for hockey. For each of the events listed above and for each location on the ice we calculated the probability that that event would lead to a goal by each team in the 20 seconds following that event. Our value is the probability that a goal will result for the home team minus the probability that a goal will result for the away team. We refer to this as the net probability after 20 seconds or NP20. Twenty seconds was chosen after an analysis of the changes in these probabilities in the seconds after each event. Changes after 20 seconds were not significant. The exceptions to this valuation of events are shots, goals and penalties. We treated both shots and goals as shots since there is strong evidence that shooting percentage regresses strongly toward the mean, e.g. Desjardins [10]. To assess the value of a shot or a goal, we take the NP20 for the shot and add to it the probability that the shot would be a goal. The probability that a shot would be a goal was broken down by RTSS shot type (wrist, slap, snap, backhand, tip-in, deflection and wraparound) as well as gridded shot location (based upon the adjusted x- and y- coordinates). For each shot we broke the offensive zone into 54 grids based upon the adjusted shot location. We then calculated the probability that a shot would be a goal for each grid. For penalties we multiplied the length of the penalty in minutes times the league average power play success rate per minute to determine the value for a penalty. Note that values are negative relative to the team committing the penalty. Below we will refer to the values of all events as NP20. It is worth noting here that the values for events will be positive for events that benefit the home team and negative for events that benefit the away team. Appendix II has some example NP20 values for a variety of events.

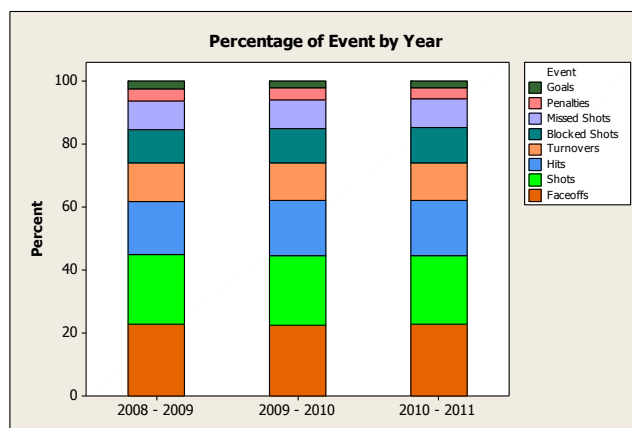


Figure 1: Percentage breakdown of all NHL RTSS events per season

Finally in this section, we note the predictability and the repeatability of events and shots in the NHL. There are approximately 300,000 of the events listed above recorded by the NHL's RTSS system for each regular season. Figures 1 and 2 below give the breakdown of the percentage of each RTSS event (Figure 1) and each RTSS shot type over three NHL regular seasons. As we can see from these graph, from one season to the next there is a great deal of similarities in the percentages for each play and the proportion of shots of a given type.

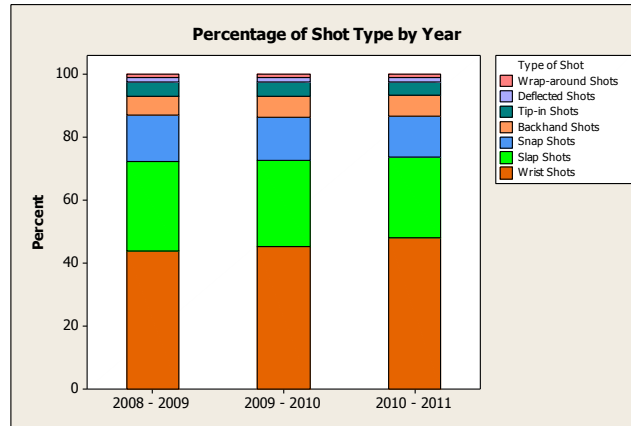


Figure 2: Percentage breakdown of all NHL Shots by RTSS shot type and season

### 3 Methodology

The goal of our methodology here is to create a ratings system for NHL forwards and defensemen that values their role in creating goals as well as preventing them. Having collected and processed the data described in the previous section, we use a model with the values for each play as the response. Our model is the following:

$$NP20 = \mu + \sum_{j=1}^P 1_{ij}^H \theta_j + \sum_{j=1}^P 1_{ij}^A \theta_j + \gamma ZS, \quad (1)$$

where  $\mu$  is the impact of home ice advantage on each play,  $\theta_j$  is the effect of player  $j$ ,  $P$  is the total number of players who have been on the ice of at least one event and  $\gamma$  is the effect of a zone start on the NP20 of each event. The variables  $1_{ij}^H$  and  $1_{ij}^A$  are defined to be

$$1_{ij}^H = \begin{cases} 1 & \text{if player } j \text{ is on the ice for the home team,} \\ 0 & \text{otherwise,} \end{cases}$$

and

$$1_{ij}^A = \begin{cases} 1 & \text{if player } j \text{ is on the ice for the away team,} \\ 0 & \text{otherwise.} \end{cases}$$

We fit this model using ridge regression following recent work in hockey as well as other sports, Macdonald [1] and [11]. One outcome of using ridge regression is that players with smaller sample sizes have their ratings deflated. This accounts for larger fluctuation in the ratings of players with smaller sample sizes, i.e. fewer number of events for which they were on the ice. Another use of ridge regression is to account for multicollinearity, the correlation of predictors in a regression. In hockey there are many players that are often on the ice together (e.g. Henrik and Daniel Sedin of the Vancouver Canucks) and ridge regression is useful to deal with this form of multicollinearity. Our ridge parameter is then chosen to minimize the predictive error in our ratings. In our data we treated players as different if they played on two different teams. So, for example, Michael Cammalleri is treated as two different players, one when he played for the Montreal Canadiens and one when we played for the Calgary Flames after being traded in the middle of the 2011-12 regular season. In doing this we can look at the variability between the ratings when a player is on one team and when they are on another. For our ratings to have maximal utility and if our ratings are capturing true isolated player performance, we want those differences to as small as possible. To that end we created the following metric:

$$\Gamma = \frac{\frac{1}{N_T} \sum_{k=1}^T n_k (\hat{\theta}_{k1} - \hat{\theta}_{k2})^2}{\frac{1}{N} \sum_{j=1}^P n_j (\hat{\theta}_j - \bar{\theta})^2}.$$

This ratio represents the variability in ratings of players who are traded to the variability in all players. In the above equation, the numerator consists of  $\hat{\theta}_{k1}$  and  $\hat{\theta}_{k2}$  which are the ratings of the  $k^{\text{th}}$  player for their first and second team, respectively,  $n_k$  which is the average number of players in which the  $k^{\text{th}}$  player was involved and  $N_T$  which is the sum of all  $n_k$ 's.  $\hat{\theta}_j$  is the estimated rating of player  $j$ ,  $\bar{\theta}$  is the average rating of all players,  $n_j$  is the number of plays in which player  $j$  was involved, and  $N$  is the sum of all  $n_j$ 's.  $T$  represents the total number of players who were traded. For a reliable ratings system, we want  $\Gamma$  to be small. We used an iterative process to guide our selection of the best ridge regression parameter for  $\Gamma$ .

Our approach has several advantages. First, we are evaluating players for every event that happens while they are on the ice. While in the short term there is the potential for chance or unlikely events to influence a player's performance, we use two regular season's worth of data for our ratings to isolate the individual effect of a given player. Second, we are adjusting for the quality of other players on the ice both those playing with and those playing against a given player as well as where a player's shift starts. These factors are well known to impact performance and so they explicitly part of our model. Third, the use of NP20 as a response maximizes the information available from the NHL's RTSS system. Given the low scoring nature of hockey, we probabilistically value each play by the impact it has on goal scoring. Fourth, our approach looks at events relative to what we would expect and so the THoR ratings reward players who are good at both ends of the ice. If on average a player is involved in more events that lead to goals then they will have a higher THoR. Finally, as we will see below our approach is robust in that player value is consistent when players change teams.

## 4 Results

We applied the above methodology to all even strength events from the 2010-11 and 2011-12 regular seasons. There were over 300,000 events in each season. As mentioned above, our focus was on even strength data because most NHL players have a significant amount of time spent at even strength. We will discuss a possible extension of our model for powerplay and penalty kill situations below. After fitting our model and obtaining a per event rating for each player, we multiplied that rating by  $80 \times 82/6$  to get our THoR values. The value 80 was chosen since this is approximately the number of events at even strength per game that a typical player is involved in. We next multiplied this value by 82, the number of games played by each team in a *non-lockout* NHL season. Thus, we get a rating that does not depend upon the number of games or events for which a player was on the ice. In doing this multiplication, we get values for THoR that are the number of goals both for and against that all players would contribute if they all played the same number of games relative to an average player. Consequently, our ratings are relative to an average player. That is, an NHL player for whom the events on the ice are at the league average adjusting for the other factors in the model will have a THoR of zero. It should be noted that the use of ridge regression means that we are 'shrinking' player ratings relative to ordinary regression. This has the effect of moving players with smaller number of events toward a rating of zero. Here, we chose a scaled ridge parameter which resulted in a  $\Gamma$  value of 0.12. This values means that the average variability in players who change teams is 12% of the variability in ratings of all players.

In addition to player ratings, we also gained estimates of the home ice effect and of the zone start effect. Our per play estimate of home-ice advantage is approximately 0.32 goals per game. Note that for the two years we are considering here Sagarin's team level ratings of home ice advantage were 0.16 and 0.58 goals per game, respectively, Sagarin [12, 13] which gives an average of 0.37. Buttrey et al [14] found a home ice advantage of approximately 0.21 for the 2008-9 NHL regular season. Thus, our estimate of home ice advantage seems appropriate. The effect of zone starts is much stronger than the effect of home ice on a per event basis. We find that starting in a team's offensive zone is the equivalent of replacing an average player with one of the top five forwards in all of hockey. On average, this amounts to adding an extra 0.53 goals per game. That is, for a player starting all of their shifts in a given game in the offensive zone, we expect that that player will produce an additional half goal per game. For a player that starts just 10 additional shifts in the offensive zone per game, we estimate that they will create an additional goal differential of 5.4 goals over the course of an 82-game season.

Tables 1 and 2 present the top rated players based upon the THoR methodology for Defensemen and Forwards, respectively. Players in those tables had to appear in 4000 events over two seasons, which is the equivalent to being in approximately 47 (out of 162 possible) games over two seasons. Those ratings are in wins which comes from taking the goals created over the course of a season (THoR ratings  $\times 80 \times 82$ ) and dividing by 6, Vollman [15]. The THoR top defenseman is Kimmo Timonen of the Philadelphia Flyers. On average over the two most recent NHL seasons, Kimmo was responsible for almost six wins per season for the Flyers. Timonen was followed by Drew Doughty, Tom Gilbert, Fedor Tyutin and Mark Giordano in the top five of all defensemen. A THoR rating of 4.07 for Drew Doughty means that he was worth just over 4 more wins per year than the average player over the last two years.

Table 1: Wins Created by Top 15 Defensemen

Team	Player	Position	Wins Created
Philadelphia Flyers	Kimmo Timonen	D	5.73
Los Angeles Kings	Drew Doughty	D	4.07
Edmonton Oilers	Tom Gilbert*	D	3.32
Columbus Blue Jackets	Fedor Tyutin	D	3.13
Calgary Flames	Mark Giordano	D	3.08
Philadelphia Flyers	Andrej Meszaros	D	2.82
Chicago Blackhawks	Brent Seabrook	D	2.63
New York Rangers	Ryan McDonagh	D	2.50
Detroit Red Wings	Niklas Kronwall	D	2.48
Anaheim Ducks	Lubomir Visnovsky*	D	2.48
Pittsburgh Penguins	Paul Martin	D	2.27
Winnipeg Jets	Tobias Enstrom	D	2.23
Ottawa Senators	Erik Karlsson	D	2.22
Boston Bruins	Zdeno Chara	D	2.18
New York Rangers	Michael Sauer	D	1.95

\*Tom Gilbert is now with the Minnesota Wild, Lubomir Visnovsky's rights were recently traded to the New York Islanders

In Table 2, we have the top 15 NHL forwards. Alexander Steen has been the top performing forward creating approximately 6.72 wins per season for the St. Louis Blues relative to an average player. The next best two-way forward is Pavel Datsyuk. After these two players, the top players in terms of wins created have been Tyler Kennedy, Patrice Bergeron and Patric Hornqvist. Certainly among that group Tyler Kennedy is the biggest surprise. Playing for the Pittsburgh Penguins, he is often on their 3<sup>rd</sup> line and is generally not considered an elite player; however, he has averaged 39 points over the last two seasons while playing against above average competition and with below average teammates. We also note that had Sidney Crosby, who is ranked 15<sup>th</sup>, played two full seasons he would have been in the top 10 and possibly in the top 5. He appears lower here due to the shrinkage of the ridge regression for players with smaller sample sizes.

Table 2: Wins Created by Top 15 Forwards

Team	Player	Position	Wins Created
St. Louis Blues	Alexander Steen	C	6.72
Detroit Red Wings	Pavel Datsyuk	C	6.32
Pittsburgh Penguins	Tyler Kennedy	C	6.05
Boston Bruins	Patrice Bergeron	C	5.95
Nashville Predators	Patric Hornqvist	R	5.88
Phoenix Coyotes	Ray Whitney <sup>+</sup>	L	5.62
Pittsburgh Penguins	Evgeni Malkin	C	5.57
Vancouver Canucks	Ryan Kesler	C	5.53
Chicago Blackhawks	Jonathan Toews	C	5.50
Vancouver Canucks	Daniel Sedin	L	5.47
San Jose Sharks	Joe Pavelski	C	5.42
Toronto Maple Leafs	Mikhail Grabovski	C	5.13
Carolina Hurricanes	Jeff Skinner	C	5.07
Los Angeles Kings	Anze Kopitar	C	4.93
Pittsburgh Penguins	Sidney Crosby	C	4.92

<sup>+</sup> Ray Whitney has signed as a free agent with the Dallas Stars



Comparing forwards and defensemen on THoR, we can see that the best forwards have higher ratings than the best defensemen. Kimmo Timonen, our highest rated defensemen, is the 6<sup>th</sup> rated player overall and Drew Doughty the second best defensemen would be the 30<sup>th</sup> rated THoR player. This relationship is also true if we consider the average of the two positions. The average THoR for forwards is 2.76 with a standard deviation of 9.88 and the average THoR for defensemen is -1.76 with a standard deviation of 9.34. This suggests that for even strength that forwards are more important for creating goals than defensemen. Note that on average defensemen tend to be on the ice more than forwards. Appendix III has a list of the top 50 NHL players based upon THoR.

## 5 Conclusions

In this paper we have proposed a new statistical methodology, the total hockey rating (THoR), for the two-way performance evaluation of National Hockey League forwards and defensemen. This approach is based upon every even strength event that happens while a player is on the ice and accounting for which team has home ice, where the player's shift started (zone starts) and the other players on the ice both on that player's team and on the other team. As with other analyses we have focused here on even strength since almost all NHL players play extensive even-strength minutes and this ensures that we have sufficient sample size for this analysis. Our approach here is a probabilistic one necessitated by low NHL scoring rates. We value each event by the net probability it will lead to a goal for the home team minus that same probability for the away team. To obtain our ratings, we fit our model using ridge regression to all of these events for two NHL regular seasons.

In order to create THoR, we developed a probabilistic methodology that assigns value to each action event that is recorded by the NHL as part of their RTSS system. To overcome some of the limitations of the RTSS system we created new adjustments for shot locations and we treated both takeaways and giveaways as turnovers. Further, we introduced a new metric,  $\Gamma$ , for the evaluation of a rating system that looks at the ratio of variability in ratings for players that change teams to the overall variability in all player ratings. Our resulting ratings are the effect of an individual player on each event when they are on the ice. To get the Total Hockey Ratings (THoR), we turn this per event value into a per season estimate of the number of wins created. THoR values are then directly comparable as they are based upon treating each player as if they played the same amount of time.

We have presented the results of THoR for the top 15 Forwards (Centers and Defensemen) and the top 15 Defensemen based upon two complete NHL regular season's worth of even strength data. A complete list of the top 50 NHL players is found in Appendix III. Only three defensemen appear among the top 50 THoR rated players. Further, defensemen are found on average to not be as valuable at even strength as forwards. One possible reason for this is that the value of defensemen is less pronounced at even strength. In order to understand this phenomenon, we plan to extend the THoR model to powerplay situations.

Based upon our analysis the top players are worth over five wins per season for their respective teams. The THoR ratings evaluate the number of additional wins over an average player that can be attributed to a given player. These wins are measured in two-way contribution. That is, they are based upon the expected number of goals created and prevented. Thus, THoR gives a complete or *total* evaluation of a given player.

## 6 Acknowledgements

Some of the groundwork for this analysis was begun by Dennis F Lock (graduate student in Statistics at Iowa State University) and Matt Generous (currently playing for Rauman Lukko in SM-Liiga, the Finnish Elite Hockey League). Ed Harcourt from St. Lawrence University assisted in the development of the code for these analyses. This material is based upon work supported by the National Science Foundation under Grant No. 0959713.

## 7 References

- [1] Brian Macdonald. "Adjusted Plus-Minus for NHL Players using Ridge Regression with Goals, Shots, Fenwick, and Corsi." Web. 1 Jan 2012 <http://arxiv.org/pdf/1201.0317.pdf>.
- [2] Vic Ferrari. "Shots Directed At Net." Web. 5 Mar 2006 <http://vhockey.blogspot.com/2006/03/shots-directed-at-net.html>.
- [3] Tom Awad. "Numbers on Ice: Understanding GVT, Part 1." Web. 30 Jul 2009 <http://hockeyprospectus.com/article.php?articleid=233>
- [4] Alan Ryder. "Product Recall Notice for Shot Quality." Web. 1 Jun 2007 <http://hockeyanalytics.com/2007/06/product-recall-notice-for-shot-quality/>.
- [5] Gabriel Desjardins (Hawerchuk). "Giveaways and Takeaways." Web. 12 Oct 2009 <http://www.arcticicehockey.com/2009/10/12/1081096/giveaways-and-takeaways>.
- [6] John Fischer. "Blocked Shots, The New Jersey Devils' Rink, and Scorer Bias - A Follow Up." Web. 8 Jul 2010 <http://www.inlouwetrust.com/2010/7/8/1559914/blocked-shots-the-new-jersey>.
- [7] M. E. Schuckers, "DIGR: A Defense Independent Rating of NHL Goaltenders using Spatially Smoothed Save Percentage Maps" Proceedings of the MIT Sloan Sports Analytics Conference, March 2011.
- [8] Cam Charran. "The Effect Of Zone Starts On Offensive Production." Web. 22 May 2012 <http://nhlnumbers.com/2012/5/22/the-effect-of-zone-starts-on-offensive-production>
- [9] Alexander Calloway. "More NHL Metrics: A Beginner's Guide to Zone Starts." Web. 27 Aug 2012 <http://www.litterboxcats.com/2012/8/27/3268425/a-beginners-guide-to-zone-starts-nhl-advanced-metrics-panthers-florida>
- [10] Gabriel Desjardins (Hawerchuk). "Luck vs Shot Quality in Shooting Percentage." Web. 25 Oct 2011 <http://www.arcticicehockey.com/2011/10/25/2512376/luck-vs-shot-quality-in-shooting-percentage>
- [11] "An Advanced Stats Primer for the NBA." Web. 6 Dec 2011 <http://www.goldenstateofmind.com/2011/12/6/2602153/advanced-stats-primer>
- [12] Jeff Sagarin. "Jeff Sagarin NHL ratings." Web. 11 Jun 2012 <http://usatoday30.usatoday.com/sports/sagarin/nhl1112.htm>
- [13] Jeff Sagarin. "Jeff Sagarin NHL ratings." Web. 15 Jun 2011 <http://usatoday30.usatoday.com/sports/sagarin/nhl1011.htm>
- [14] Buttrey, Samuel E et al. "Estimating NHL Scoring Rates," *Journal of Quantitative Analysis in Sports*, vol 7, Issue 3, July 2011.
- [15] Rob Vollman "Howe and Why: Goals Versus Salary, 2011-12." Web. 10 Oct 2012 <http://www.hockeyprospectus.com/article.php?articleid=1393>.

## 8 Appendix I: Shot Coordinate Adjustments

Our adjustment is based upon the following equations:

$$\begin{aligned}x' &= F_X^{-1}(F_R(x) - (F_{RA}(x) - F_A(x))), \\y' &= G_Y^{-1}(G_R(y) - (G_{RA}(y) - G_A(y))),\end{aligned}$$

where  $x$  and  $y$  are the original locations of a given shot,  $F_X$  and  $G_Y$  are the cumulative distribution function (cdf) for the  $x$ -coordinates and the  $y$ -coordinates, respectively. The adjust values for  $x$  and  $y$  are  $x'$  and  $y'$ , respectively. Then  $F_R$  and  $G_R$  are the cdf's for  $x$  and  $y$  coordinates for rink  $R$  at which a given shot was taken, while  $F_{RA}$  and  $G_{RA}$  are the cdf's for  $x$  and  $y$  for all shots taken by the away team at rink  $R$ . Finally,  $F_A$  and  $G_A$  are the cdf's of  $x$  and  $y$  coordinates for all away shots. We start by finding the cumulative probability for a given shot at rink  $R$ , say  $F_R(x)$ , we then adjust this probability by how different shots by the away team at that location are from all shots by the away team,  $F_{RA}(x) - F_A(x)$ . After this difference is subtracted we invert the adjusted probability to the original scale to determine our adjusted value,  $x'$ . As part of this process we have maintained the discrete nature of the  $x$ - and  $y$ - coordinates and consequently,  $x'$  and  $y'$  are measured in the same whole units (feet) as the original  $x$  and  $y$  measurements. One possible future extension would be to have  $x'$  be based upon a smoothed version of  $F_X^{-1}$ .



## 9 Appendix II: NP20 Values for Selected Events

In the table below we present some selected values for NP20 given different on-ice events, their location and, if relevant, the shot type. For all of these values the team carrying out the event is the Home team. That is, the home team is taking the shot, gaining a turnover or hitting an opponent. The equivalents for the Away team are approximately the negative of the values given here.

Event	Shot Type (if relevant)	Location	NP20
SHOT	Backhand	Off	0.1348
SHOT	Wrist	Off	0.1096
SHOT	Slap	Off	0.0697
TURN(gained by Home Team)		Off	0.0362
FAC		Off	0.0167
MISS	Wrist	Off	0.0159
HIT		Off	0.0039
FAC		Neu	0.0026
HIT		Neu	-0.0008
TURN (gained by Home Team)		Neu	0.0264
FAC		Def	0.0005
HIT		Def	-0.0060

## 10 Appendix III: Top 50 NHL Forwards (F) and Defensemen(D)

Rank	Team	Player	Position	THoR Wins Created
1	St. Louis Blues	Alexander Steen	F	6.72
2	Detroit Red Wings	Pavel Datsyuk	F	6.32
3	Pittsburgh Penguins	Tyler Kennedy	F	6.05
4	Boston Bruins	Patrice Bergeron	F	5.95
5	Nashville Predators	Patric Hornqvist	F	5.88
6	Philadelphia Flyers	Kimmo Timonen	D	5.73
7	Phoenix Coyotes	Ray Whitney	F	5.62
8	Pittsburgh Penguins	Evgeni Malkin	F	5.57
9	Vancouver Canucks	Ryan Kesler	F	5.53
10	Chicago Blackhawks	Jonathan Toews	F	5.50
11	Vancouver Canucks	Daniel Sedin	F	5.47
12	San Jose Sharks	Joe Pavelski	F	5.42
13	Toronto Maple Leafs	Mikhail Grabovski	F	5.13
14	Carolina Hurricanes	Jeff Skinner	F	5.07
15	Los Angeles Kings	Anze Kopitar	F	4.93
16	Pittsburgh Penguins	Sidney Crosby	F	4.92
17	Buffalo Sabres	Jason Pominville	F	4.78
18	Carolina Hurricanes	Eric Staal	F	4.53
19	Colorado Avalanche	Matt Duchene	F	4.42
20	Los Angeles Kings	Ryan Smyth	F	4.33
21	New Jersey Devils	Patrik Elias	F	4.32
22	Detroit Red Wings	Henrik Zetterberg	F	4.30
23	Edmonton Oilers	Taylor Hall	F	4.30
24	San Jose Sharks	Logan Couture	F	4.28
25	Colorado Avalanche	Paul Stastny	F	4.25
26	New Jersey Devils	Zach Parise	F	4.25
27	Chicago Blackhawks	Viktor Stalberg	F	4.23
28	Columbus Blue Jackets	Brandon Dubinsky	F	4.18
29	San Jose Sharks	Patrick Marleau	F	4.13
30	Los Angeles Kings	Drew Doughty	D	4.07
31	St. Louis Blues	Andy McDonald	F	3.97
32	Columbus Blue Jackets	Antoine Vermette	F	3.87
33	Chicago Blackhawks	Patrick Sharp	F	3.85
34	Calgary Flames	Tim Jackman	F	3.83
35	New York Islanders	Kyle Okposo	F	3.82
36	Nashville Predators	Martin Erat	F	3.75
37	Philadelphia Flyers	Claude Giroux	F	3.73
38	Vancouver Canucks	Henrik Sedin	F	3.70
39	Anaheim Ducks	Ryan Getzlaf	F	3.68
40	Nashville Predators	Colin Wilson	F	3.63
41	Chicago Blackhawks	Marian Hossa	F	3.62
42	San Jose Sharks	Torrey Mitchell	F	3.62
43	Anaheim Ducks	Corey Perry	F	3.45
44	St. Louis Blues	David Backes	F	3.43
45	Montreal Canadiens	Erik Cole	F	3.42
46	Ottawa Senators	Jason Spezza	F	3.35
47	Ottawa Senators	Daniel Alfredsson	F	3.35
48	Montreal Canadiens	Brian Gionta	F	3.33
49	Washington Capitals	Alexander Semin	F	3.33
50	Minnesota Wild	Tom Gilbert	D	3.32