A Comparison of Drafting Performance between the NHL's Central Scouting Service and NHL teams

Michael Schuckers[1]
St. Lawrence University, and Statistical Sports Consulting, LLC
michael.schuckers@statsportsconsulting.com

Abstract
In this paper we evaluate the performance of National Hockey League (NHL) teams in the entry draft to see if they generate risk-adjusted return in excess of the market index. In place of an index, we use the annual rankings of the NHL's Central Scouting Service (CSS) on the selection of NHL Draft picks. The analysis we present below is a statistical one based upon the 1998 through 2002 NHL entry drafts. Several authors have previously written statistical analyses of the NHL draft. Most of these analyses have focused on the value of an NHL draft pick or the likelihood that a player would eventually play a certain number of games. In this paper we focus on the added benefit that teams get from their own internal scouting by comparing the team's actual selections to those suggested by the CSS's final rankings. We find that the additional information that teams have makes them significantly better than the CSS at selecting the optimal remaining player at a given position. Further, we quantify the monetary return on team scouting by assessing how much better teams did than Central Scouting and putting a value on that additional performance. We estimate that teams, on average, get a return of between 4 to 10 million $US annually from their scouting; however, we also find that no team significantly outperformed the others in selecting players.

Introduction and Data

With the 22nd pick in the 2008 National Hockey League (NHL) Entry Draft, the Edmonton Oilers selected Jordan Eberle, center. With the 25th pick the Calgary Flames selected Greg Nemisz, center. Since they were drafted Eberle has played 195 NHL games, while Nemisz has played in 15 NHL games. Going into the draft the NHL's Central Scouting Service (CSS) had ranked Eberle as the 33rd best North American skater while Nemisz was ranked 22nd best in that same category. The Oilers had additional information from their internal scouting staff about Eberle and Nemisz as well as about other centers that were available with the 22nd pick including Daulten Leveille, Derek Stepan and Tyler Ennis. Each NHL team had available to them information about possible draftees from the CSS as well as information compiled by their own staffs. The outcomes of draft decisions like the ones that Edmonton and Calgary made are full of variation. Over half of all players selected in the NHL Entry Draft never play a single game in the NHL. Predicting how an 18 year old will develop is an imperfect science at best. In this paper, we will look at the quality of rankings by the CSS and draft order by team in order to evaluate and quantify the value of the additional information that teams have. In doing so, we will estimate the return that teams get from their internal scouting.

[1] Thanks to Steve Argeris for discussions that led to this article and Brian Chezum for discussions about the analysis.

The data that we analyzed here is from the five NHL Entry Drafts from 1998 to 2002. For each player selected we have their selection number, their position, the career games played (GP) in the NHL, the ranking by player type by CSS and their career goals versus threshold (GVT). These data were compiled from nhl.com, hockey-reference.com, and eliteprospects.com. We used the CSS final rankings released prior to the draft. These rankings were obtained via contemporaneous media accounts accessed via Internet searches and the Lexis/Nexis database. A player's position was recorded as either center (C), defensemen (D), forward (F), goalie (G), left wing (L) or right wing (R). In the analyses that follow we will categorize players as forwards (F) which includes C's, L's, and R's, defensemen (D) or goalies (G). The selection number for a player is the place in the draft order that they were selected. That is, the 10th player selected will have a selection number of 10.

CSS ranks players by category, either North American or European, and by whether they are a skater or a goalie. Thus, CSS produces four separate rankings without a correspondence between them for comparison purposes. To better utilize the CSS rankings, we employed Iain Fyffe's "Central Scouting Integratinator" (Cescin) metric, Fyffe (2011). Cescin takes the rankings of players by CSS within their given category and multiplies them by a factor based upon historical draft selection records. For players who were not ranked by CSS, we gave those players values of Cescin that was larger than the values produced by the original Cescin. Next, we took the values generated by Cescin and ranked them to produce the CSS orderings we use in the rest of the paper. Two response metrics are considered below: GP and GVT. Goals Versus Threshold (GVT) is a metric of player value created by Tom Awad that allocates value in team performance among the individuals on a given team, Awad (2009). The units for GVT are goals so that a GVT of 7.5 credits a player with producing 7.5 goals over replacement level. Since players can have negative GVT, we give players who never played in the NHL a value of GVT below the lowest value in our database. Our focus is on the top 210 selections since that is the length, over seven rounds, of the five most recent NHL Entry Drafts (2008 to 2012). Thus, for each year we have 210 observations except for 2002 when the 123rd pick in the draft was invalidated since the Edmonton Oilers selected a player who was ineligible to be drafted.

From analyses like Schuckers (2011) which focused on GP as the response, it is clear that for a given selection, average performance for all subsequent selection decreases. For example, the average player taken with selection 12 is better than the average of all players taken at selection numbers 13 and higher. There are and will continue to be exceptions to this generalization since this is a stochastic statement about averages rather than individuals. Many of the critiques of team drafting are that for a given selection a player whose future performance exceeded the current selection was often available. See, for example, Tingling (2011). Given the difficulty with projecting the future performance, it is important to focus on the trends rather than individuals. Clearly there is monotonicity in average performance of players versus draft selection and clearly there are long tails to the distributions of player performance for a given selection.

For the data that we are considering, we had 1049 total players. Among these, 595 were forwards, 332 as defensemen (D's), and 122 as goalies (G's). For GP, 54% of the players selected never played a game in the NHL. Given that fact, the mean number of games played per selection was 117.5 and the standard deviation was 215 games. For GVT, there were no values of that metric for players that did not appear in an NHL game. Among those who did play at least one game the worse GVT that was calculated was -27.7, the median GVT was 1.4 and the mean was 14.8. For completeness and comparability, we chose to give those who did not appear in an NHL game a GVT of negative 50, -50. We selected this value to be substantially below the other vales in our data and to permit a more complete analysis. To gauge team preferences for players we will use the selection at which the player was taken. For assessing how CSS ranked players we will use Cescin values.

There has been a good deal of previous work on evaluation of NHL draft picks. Much of this work has focused on the value of an individual selection. Johnson(2006), Tango(2007), Awad(2009), Gregor (2011) and Schuckers (2011) have all looked at methods for evaluating what an individual draft pick is worth in terms of a measure of value. Below we will use Career GVT and Career GP. Doing so slightly weights earlier drafts (e.g. 1998) more heavily than later drafts (e.g. 2002) in our analysis. We compensate for this by using statistical methods that are more robust to tail behavior. Recently, Eric Tulsky wrote about the value of a pick in terms of the trade market, Tulsky(2013). This analysis more closely resembles the original intent of the National Football League's Value Pick Chart, a quick reference developed by then-Dallas Cowboys coach Jimmy Johnson in the late 1980s that came into widespread acceptance over the next decade, (McGuire 2010).

The focus of our analyses is the value that teams get from their scouting departments. It is difficult to quantify the value of CSS since teams have information both from CSS and from their own internal scouting departments. We propose that the CSS represents a suitable, if crude, proxy for a benchmark index. Teams could, theoretically, get rid of their amateur scouting department and rely on CSS's rankings, as it typically ranks more than the 211 players drafted.[2] Below we will look at the difference in how teams rated players and how CSS rated players to get an idea about the value added by team scouting staffs. The number of scouts employed by teams varies considerably. For example the New York Islanders have 11 individuals with scouting responsibilities while the Toronto Maple Leafs list 23 on their respective webpages. Another small market U.S.-based team estimates that they spend approximately $2 million on their annual scouting budget.

In the rest of the paper we begin by considering how often CSS rankings and team draft order were able to optimally or nearly optimally selected the best player at a given selection. We find that teams outperform the CSS rankings. Next we consider a non-parametric regression following Schuckers (2011) of our performance metrics onto the player orderings, both CSS and team. This approach also finds that teams outperform the CSS. Finally, we consider a novel approach that looks at the relationship in the rank orderings and relative value of our performance metrics. Since this final approach is conditional at the individual level, it is the most relevant and informative. We then calculated the excess value above the CSS rankings that teams get in terms of GP and GVT and these are roughly $6 million dollars per year.

[2] Though we note that this would be rather inadvisable for strategic reasons, including but not limited to teams would have little idea who other teams' top prospects were for trade value.

Quality of Central Scouting Draft Order

In this section we look at how well CSS and teams drafted players for a specific position. Our first assessment is on the ability of CSS to correctly order the possible draft selection based upon CSS's own ordering of draftees. To assess this we looked at the percent of times that the ordering by CSS as reflected in Cescin resulted in the optimal ordering at a given position, either C, D, F, G, L or R. Note that this differs from the way the CSS ranks players which combines centers, forwards, defensemen, left wings and right wings into skaters. We also considered the percent of times that the CSS came within approximately one-half standard deviation of the optimal choice at a given position. We determined if a selection was optimal by considering all of the remaining draftees in a given year at the same position as the selected player. If selected player had the highest metric among all other available players, then that player was considered the optimal selection. For GP, the standard deviation (SD) of the players taken in the first 210 selection was 215 games, while the standard deviation of those same players for GVT was 21. Table 1 below has the results of this analysis for GP and GVT. Overall it is clear that team ordering (based upon the actual draft) is better than CSS ordering. Simply choosing the best available player happens about 6% of the time using CSS and about 10% of the time with team ordering. Ordering players so that the current selection is within ½ SD of the best of the lower ranked players happens about 11% of the time for Central Scouting. For teams this latter value is 26% for GP and 15% for GVT.

Table 1: Comparison of Performance of CSS and Actual Draft Orderings

Metric	Ordering	Percent of draft picks where ordering was optimal	Percent of draft picks where ordering was within ½ SD of optimal
GP			
	CSS	5%	12%
	Team	11%	26%
GVT			
	CSS	7%	10%
	Team	10%	15%

Comparison of Average Performance by Player Ordering

Above, we have concentrated on the optimal or nearly optimal decision at a given selection. We next looked at the impact of these selection criteria on the average outcome variables per draft selection. To evaluate this impact we looked at the relationship between GP and Ordering and between GVT and Ordering. As before, ordering for Central Scouting is done based upon CESCIN while ordering for Teams is from the actual draft selections. To estimate this relationship we use LOESS regression as was done in Schuckers (2011). LOESS regression is a non-parametric approach that does not specify a particular form (e.g. linear, quadratic) for the relationship between the variables. Effectively it is a methodology for locally smoothing the response (GP or GVT) at each value of the predictor (selection).

In the lefthand graph of Figure 1 we have the LOESS regressions for predicting GP based upon the draft selection order from CSS(blue) and from the actual draft (red). Our expected or predicted values from

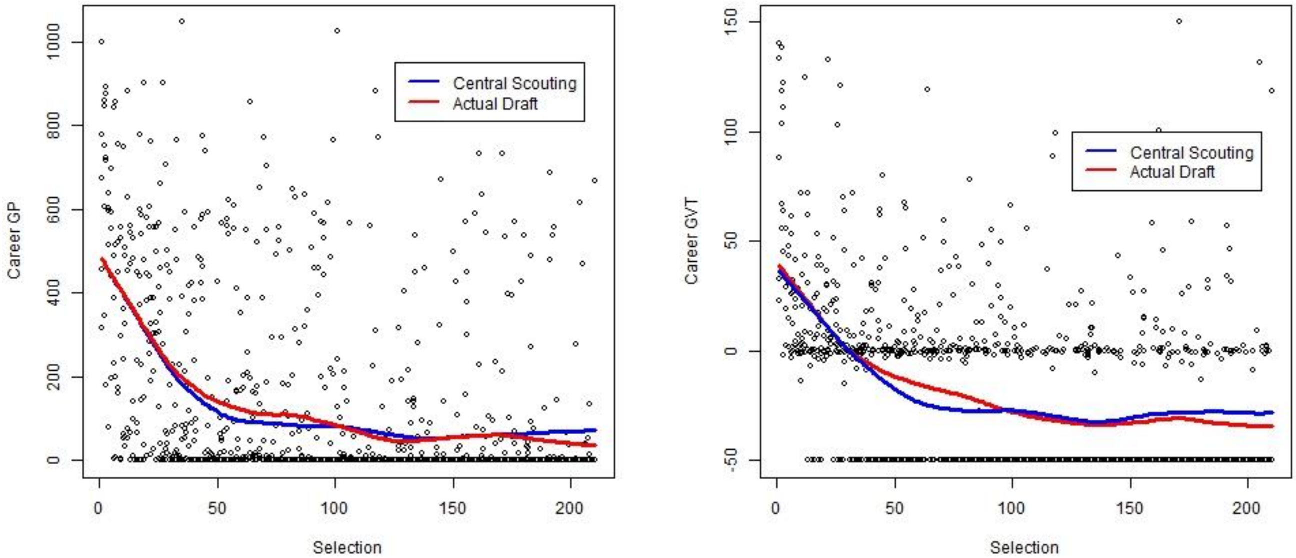


Figure 1: Average Performance Comparison for Career GP and Career GVT

both of these curves are very similar. Both start at roughly the same value for the first selections and decrease steeply until approximately the 40th selection, then less steeply until the 120th selection after which they are roughly flat. The difference between the two curves is that the draft selections outperform the CSS rankings from about 40 to 100. Currently this corresponds to roughly the early second round to the end of the third round. These are the locations where, on average, team selections are better than the ordering from Central Scouting. Central Scouting does better than teams over the last 35 selections. Turning to GVT, we see very similar results to those we found for GP. The shape of the prediction curve is strikingly similar for GVT though not as steep over the initial 50 rankings.

Rank Differential Comparison

While the above analysis gives a sense of how CSS and team scouting group perform in their rankings on average, those comparisons and rankings do not condition on some important factors such as the individual player and their position. To account for these we next consider an analysis that looks at the differential between each player’s actual selection and their CSS ranking. Below we will refer to this as rank differential. For example, Rico Fata was the 6th overall selection by the Calgary Flames in the 1998 draft. Cescin has Fata as the 13th ranked player. Consequently our rank differential for Fata would be -7 meaning he was taken seven places ahead of where CSS ranked him. Players with negative rank differential were taken earlier than CSS ranked them and players with positive rank differential were taken later than CSS ranked them. 56% of selections were rank differential positive, 43% were negative and 1% were zero. If team scouting does well then we should expect that players who have negative rank differential will also outperform what we would expect based upon CSS ranking of them and vice versa for players with negative rank differential. For both GP and GVT we looked at what each player

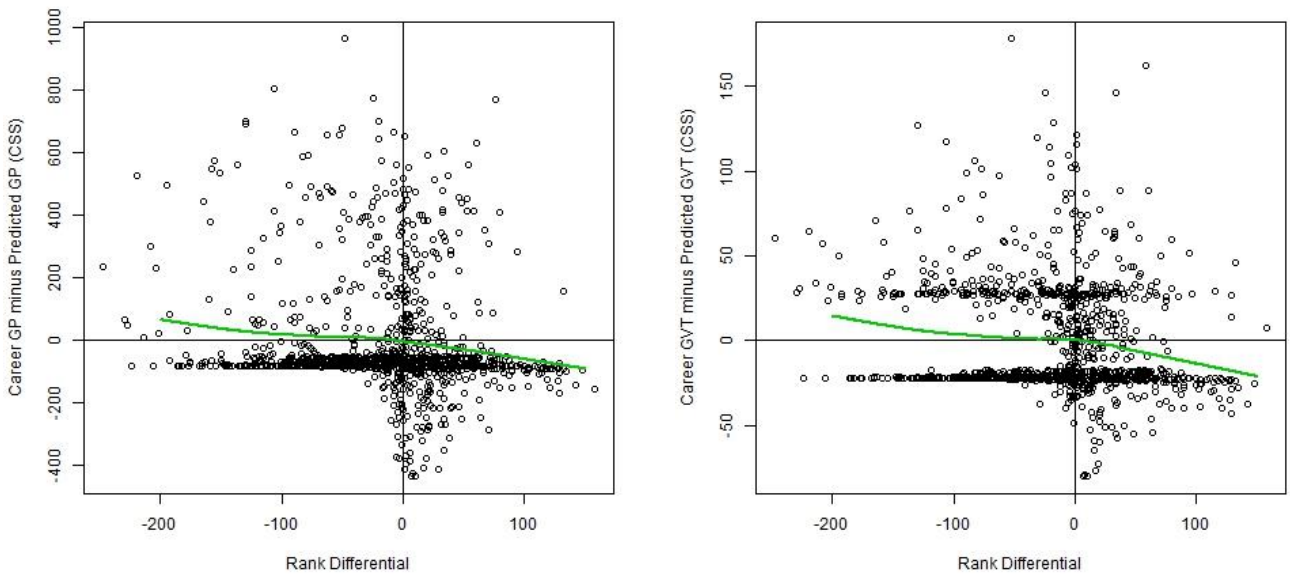


Figure 2: Relationship between Relative Performance and Rank Differential

achieved relative to what we would have expected from them based upon their CSS ranking and compared that to what we would have expected based upon their actual draft selection. To calculate the expected values we used the LOESS regressions from Figure 1. Figure 2 has a plot of the Career GP minus Expected GP for each player plotted against their rank differential as well as the same plot for GVT. If team scouting was perfect, there would only be players in the upper left and lower right quadrants of these graphs. Players in the upper left are those that exceeded CSS expectations and were taken earlier than CSS had them ranked. Players in the lower left are those that underperformed CSS expectations and teams took them earlier than CSS had them ranked.

Each graph in Figure 2 also contains a green curve that is the smoothed average difference in GP and in GVT, respectively, relative to the rank differential. It is important to note that both curves go from the upper left quadrant to the lower right quadrant and roughly pass through the origin (0,0). This suggests that when teams differentiate from the CSS ordering they are gaining some value both in terms of GP and in terms of GVT. To estimate the average net values that team scouting contributes above and beyond the CSS, we estimated the GP gained and the GVT gained at each rank differential for all players. We then scaled the total values of each to seven selections per season which is the average number of picks that a team has in the current NHL Entry Draft. For GP this yields approximately 124 games per draft, while GVT yielded about 24 goals. Since the average NHL salary per game is approximately $29300 (Dorish, 2011), we can say that in terms of GP team scouting adds approximately $3.6 million with each draft. Using GVT, we estimate that team scouting gains a team about $7.9 million in value using the metric that a goal in the NHL is worth approximately $1/3 million (Vollman, 2012).

We also analyzed the value gained over these drafts by team. There are some winners and some losers but the average gains per team per pick across these five years did not differ significantly from what would have been expected by chance (Shapiro-Wilk test p=0.13 for GP and p=0.15 for GVT). We used an average here since Atlanta, Columbus and Minnesota entered the league between 1998 and 2000 and participated in fewer drafts than the other teams. There are no outliers here. Further, we looked at the correlation in team average return over CSS for 1998-2000 and the team average return over CSS for 2001-2002. The correlations were slightly positive, r=+0.2 and r=+0.3, but not significantly so (p=0.22 and p=0.09) for GP and GVT, respectively. No team over this period outperformed the others and no team got a significantly larger or smaller return from their internal scouting.

Rank Differential Comparison By Position

Recognizing that teams often select the best player at a particular position rather than the best player available, we next turned to analyses similar to those in the previous section for each position. As before, we have grouped players into Defensemen (D's), Forward's (F's) and Goalies (G's). The only change that we made to our approach is that the rank differential we used was that for a given position. We then found the average amount gained in terms of GP and GVT for D's, F's and G's. Those values were weighted by the average number of D's, F's and G's taken in each draft which are 2.21, 3.97 and 0.81 respectively. For this analysis we found that per player teams got more value per player from F's than D's and more value from D's than G's. In fact, we found that average gains by teams for G's were not significantly different from zero. Looking at the relationships in the graphs found in the Appendix suggests that rankings of goalies are not improved by team scouting. Incorporating all of these factors, the total value that teams get from their scouting based upon this analysis is approximately $5.9 million in GP and about $12.1 million in GVT.

Conclusions

In this paper we have looked at the impact of team scouting compared to the NHL's CSS. The metrics for performance that we have used are Games Played (GP) and Goals Versus Threshold (GVT) for a player's career. While neither of these is an ideal metric they are reasonable proxies for overall player value over the course of their career. It is also important to recognize that there are some biases in these metrics specifically for players who are still active. Among these are, for example, Vincent LeCavalier and Ryan Miller. Since each team has their own scouting staff as well as access to the assessments by the CSS, we can estimate the impact of these scouting staffs. We began by looking at the chance that the ordering by CSS and by team scouting resulted in choosing optimal or nearly optimal players. At a given selection, teams were not significantly better at picking optimal players with respect to GP or GVT but were markedly better at selecting players within a half standard deviation of optimal. When we considered the percentage of picks that were within a half standard deviation of optimal at a given selection, teams did about 10% better on both GP and GVT. There is a clear benefit here to the additional information that teams have from their scouts.

Our next focus was on the average value of a player at a given ranking. Team rankings were based upon the actual order of player selection from the draft, while CSS rankings were based upon the Cescin conversions of CSS's final rankings. In this analysis we found that teams outperformed CSS over selections 40 to 100 while CSS outperformed teams after selection 175. One possible explanation for the divergence of the two curves between selections 40 and 100 is that there is generally consensus among scouts from CSS and teams on who the top players, say the top 40, are, say the top 40. Between 40 and 100 there is value to be had in using the additional information from a team's scouts. After about selection 100, our predicted values flatten and enough noise that teams do not improve on CSS rankings, on average.

The focus of this paper was the introduction of a novel way to assess how well one ordering system compared to another. We did this by looking at the expected differential from the two rankings relative to the difference in the rankings themselves. It is clear from this analysis that individual team rankings consistently outperforms CSS rankings. This should be expected since teams use considerable resources to gain further information about each potential draftee. For GP we find that the average value that a team gets from their scouting staff is $3.6 million, while for GVT we estimate the value to be $7.9 million. This is effectively an optimization across positions. Optimizing within positions gave us a larger return per draft of $5.9 million based upon GP and $12.1 million based upon GVT. Since teams draft both for positional need and also for the best player available, it seems reasonable to state that the return that teams are getting is between $4 and $10 million dollars per draft. This is the average return that teams get on the total amount that they spend on scouting. A representative for a US-based small market team estimated that their team's scouting budget is approximately $2 million which suggests that some teams are getting a significant net benefit from their scouting groups. There is a good deal of variability in both of these estimates and as we stated above GP and GVT are not perfect metrics. However, they are reasonable metrics which have been widely used. Since players with lengthy NHL careers are outliers relative to the draft, it is likely that the values here are slightly biased downward meaning that our estimate of return on scouting is likely on the low side. While our LOESS approach discounts the impact of these outliers, our approach necessarily weights players drafted in 1998 more than those drafted in 2002 since the latter group has had fewer possible games in which to perform. Further what we have here are averages across 30 teams and there was variability between teams; however, that variability was not beyond what would have been expected by chance. This means that there is not strong enough evidence to suggest that drafting quality is different across teams, at least for the five years that we considered. A sample over a longer period of time could provide additional evidence of differences amongst teams.

This analysis could be improved by having better metrics for career player performance. The two that we have used here, GP and GVT, are reasonable stand-ins. As hockey analytics develops, utilizing more advanced methods like the Expected Goals Model (EGM), Macdonald (2012), will provide better estimates for player value and the value of player scouting once they are available for historical data. In this analysis we focused on five years of NHL Draft Entry selections. In any sort of analysis regarding professional sports drafts, there are tradeoffs to be made between completeness of careers and temporal relevance of data. Additional years of draft data would provide better estimates of player value and of the value of team scouting. We also note that while team scouting outperforms the Central Scouting Service, team scouting is far from optimal. A future analysis might look at the predictive power of analytics such as league equivalencies, DesJardins (2004), Vollman (2011), to rank players and compare results from that sort of analysis to those given here. With all that in mind, however, it is clear from this analysis that teams are getting considerable financial benefit from their internal scouting.

References

Anonymous. "The Value of a Draft Pick", March 13, 2011. http://jibblescribbits.com/2011-articles/march/the-value-of-a-draft-pick.html , accessed June 11, 2013.

Awad, T. "Understanding GVT, Part 1", July 30, 2009. http://hockeyprospectus.com/article.php?articleid=233, accessed April 1, 2013.

Awad, T. "The Top 10 Greatest Drafts Ever, Part One", June 18, 2009. http://www.puckprospectus.com/article.php?articleid=187 ,accessed June 11, 2013

Desjardins, G. "League Equivalencies", December 30, 2004. http://hockeyanalytics.com/2004/12/league-equivalencies/ ,accessed June 11, 2013.

Dorish, J. "Average Salaries in the NBA, NFL, MLB and NHL", November 12, 2011. http://sports.yahoo.com/nba/news?slug=ycn-10423863, accessed June 10, 2013.

Fyffe, I. "Evaluating Central Scouting", January 19, 2011. http://www.hockeyprospectus.com/article.php?articleid=777, accessed April 1, 2013.

Gregor, J. "The Value of a Draft Pick", February 25, 2011. http://oilersnation.com/2011/2/25/value-in-acquiring-draft-picks, accessed June 11, 2013.

Johnson, D. "Draft Schmaft: From draft pick to NHLer", June 24, 2006. http://hockeyanalysis.com/2006/06/24/draft-schmaft-from-draft-pick-to-nhler/, accessed June 11, 2013.

Macdonald, B. (2012) "An Expected Goals Model for Evaluating NHL Teams and Players", *Proceedings of 2012 MIT Sloan Sports Analytics Conference*, Boston, MA.

McGuire, M. (2010) "Matt McGuire's NFL Draftology 321: Revising the trade value chart," http://walterfootball.com/nfldraftology321.php, accessed June 13, 2013.

NHL.com (2008) "NHL Draft History" http://www.nhl.com/futures/drafthistory.html, accessed June 2011.

Tango, T. (2007). “Value of draft picks in NHL”, May 22, 2007. http://www.insidethebook.com/ee/index.php/site/comments/value_of_draft_picks_in_nhl/, accessed June 2011.

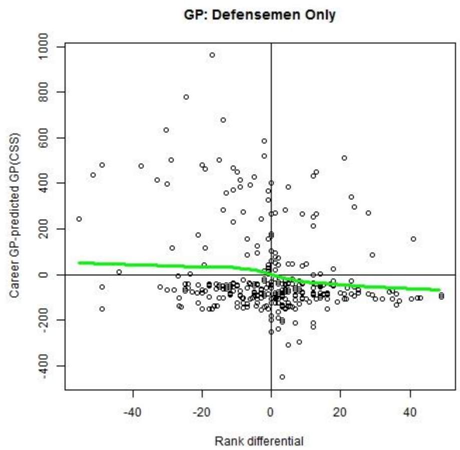
Tingling, P. et al., "Does order matter? An empirical analysis of NHL draft decisions", *Sport, Business and Management: An International Journal*, Vol. 1 Iss: 2, 2011. pp.155 – 171.

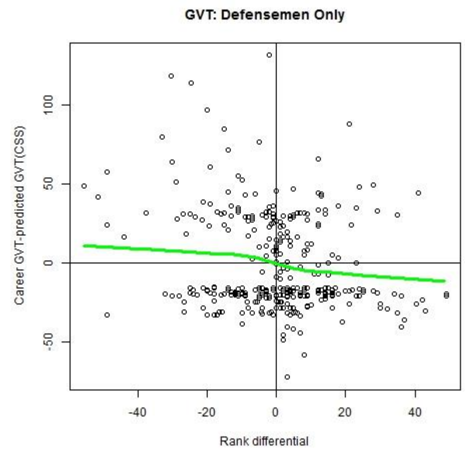
Tulsky, E. “NHL draft: What does it cost to trade up?” April 25, 2013 http://www.broadstreethockey.com/2013/4/25/4262594/nhl-draft-pick-value-trading-up, accessed June 10, 2013.

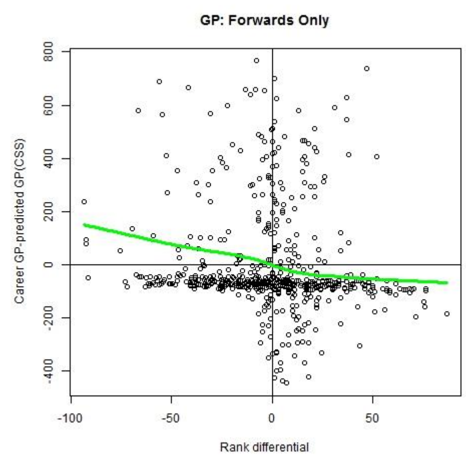
Vollman, R. “Howe and Why: AHL Translations Revisited” September 20, 2011. http://www.hockeyprospectus.com/article.php?articleid=1079, accessed June 10, 2013.

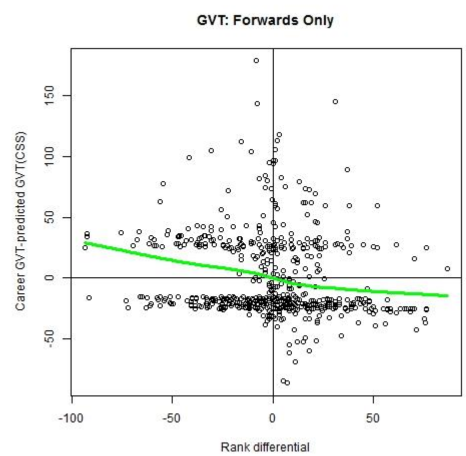
Vollman, R. “Goals Versus Salary, 2011-12”, October 10, 2012. http://www.hockeyprospectus.com/article.php? articleid=1393, accessed June 10, 2013.

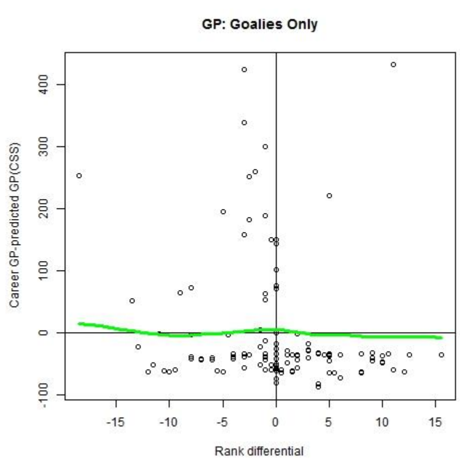
Appendix

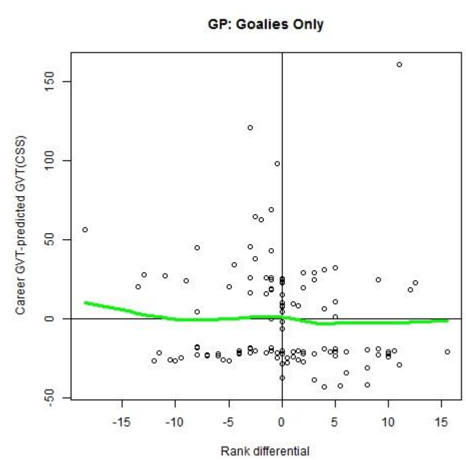

GP: Defensemen Only
Career GP-predicted GP(CSS)
Rank differential


GVT: Defensemen Only
Career GVT-predicted GVT(CSS)
Rank differential


GP: Forwards Only
Career GP-predicted GP(CSS)
Rank differential


GVT: Forwards Only
Career GVT-predicted GVT(CSS)
Rank differential


GP: Goalies Only
Career GP-predicted GP(CSS)
Rank differential


GP: Goalies Only
Career GVT-predicted GVT(CSS)
Rank differential