Using Point Shares Metrics as Descriptors of Defensive Impact in the Stanley Cup Playoffs

Adam Kiehl

July 28, 2020

Abstract

In this analysis, Hockey Reference's point shares system was used to answer the question of whether or not defenses win championships in the NHL. Previous studies conspicuously underutilized actual playoff data, so point share values were calculated based on playoff statistics alone. A logistic regression model was fit using the defensive and offensive components of point shares. It found that a small difference in defensive performance has a disproportionately large impact on the odds of a team winning the Stanley Cup. However, this is a rarity and typically, outstanding offensive performance dominates defensive play, likely due to the increased pace and aggression of playoff hockey.

1 Introduction

The phrase, 'offense sells tickets; defense wins championships,' is accredited to legendary Crimson Tide football coach Paul 'Bear' Bryant [15]. Since then, the adage has been widely circulated as wisdom by teams seeking championships through a systematic defensive stifling of their opponents. The idea is that while stakes, mistakes, tenacity, and energy all compound come playoff season, a team that can maintain a level-headed and consistent defense can capitalize on the circumstances.

This philosophy has been hotly debated by sports analysts [14] and fans [12] alike, especially when concerning football. An in-depth article on the Freakonomics blog [14] examined NFL playoff success amongst highly ranked defensive and offensive teams and found a compelling parity between the two. As of 2012, 38 Super Bowls have been won each by top-ten ranked defensive teams and top-ten ranked offensive teams. 22 and 20 have been won each by top-three ranked teams in each category. Additionally, poor defensive teams tend to lose Super Bowls at comparable rates to poor offensive teams and Super Bowls featuring a top five defense and a top five offense have been split almost exactly down the middle. The parity continues beyond the Super Bowl as well. Of the NFL playoff games to date, 62% have been won by the better offense and 58% by the better defense. Nearly the same is true of regular season games. Another study

[15] confirmed these findings by regressing seasonal yards gained and allowed on playoff wins. Yet another suggests that 'defense gets you into the playoffs but no further' and points out that recent restrictive rules imposed on defenses could tip the advantageous scales in the direction of offenses [6].

When the same questions are applied to hockey, the answers are far more scarce and less definitive. A logistic regression analysis performed by Sports Illustrated writer Ijay Palansky [16] studied the effect of the season-long goals scored differential and goals allowed differential between two Stanley Cup opponents on a team's likelihood of winning. It was found that each additional goal scored more than the opposition increased a team's likelihood of winning by 0.46% and each additional goal allowed less than the opposition increased a team's likelihood of winning by 0.37%. Redditor milkplantation [12] cited this study in a post in which he elaborated further on the correlation of other metrics with Stanley Cup success. He pointed out that between 1980 and 2018, seven Stanley Cup winners have finished first in the league in goals against and six have in goals for. Additionally, nine teams have won despite finishing tenth or worse in goals against and six have in goals for. Of the metrics of interest including shot percentage, special teams success, faceoff percentage, and PDO (the sum of a teams shot percentage and save percentage), only goaltending and shot differential seemed to correlate significantly with winning. What research that has been done appears to point to at least a parity between offensive and defensive contributions to Stanley Cup success and could suggest that an offensive-minded approach is superior.

However, what these analyses fail to account for is the rift between regular season hockey and playoff hockey. There is a widely accepted consensus between players, coaches, and fans alike that the NHL playoffs are a separate beast from the regular season. The play style is faster and more aggressive, strategies change, and the unexpected often happens. By unwritten rule, fewer penalties are called in the playoffs, creating a more physical game on both sides of the puck. In contrast with the regular season, teams play entire series in the playoffs, drastically increasing scouting and preparation time. Effort, too, apparently increases as a higher level of exertion cannot be sustained over the regular season where it often can throughout the playoffs [20]. A 2002 study [8] showed that moral reasoning in hockey players was significantly repressed in playoff situations relative to the regular season due to the higher stakes, competitive atmosphere, and unwritten codes of conduct of the playoffs. Player intensity reduces chances for passes and odd-man rushes while increasing the number of hits, blocks, and steals [5]. The hockey postseason sets itself apart from the playoffs of other sports and thus warrants a closer examination. While others have used regular season statistics to make inference about the Stanley Cup, this study will treat regular season statistics and playoff statistics as distinct, using only data from the playoffs in analysis.

1.1 Point Shares System

Hockey is also distinct from football in the structure of offensive and defensive play it presents. In football the offense and the defense are separate squads that alternate on-field play, but in hockey both forwards and defensemen are on the ice simultaneously and either can contribute both offensively and defensively. There is a much more obvious translation between a defensive play on the puck and an offensive opportunity in hockey where play is largely constant. Additionally, productive offensive teams generate opportunities, shots, and goals that all consume time in the offensive zone of the rink, minimizing the opponent's offensive zone time and lessening strain on the defense [22]. This presented the problem of quantitatively separating defense and offensive performance. Justin Kubatko, former Vice President of Sports Reference, offers a metric he developed called 'point shares' as a solution [9]. The system is based on Bill James' 'win shares' system for baseball in which players are credited with a share of its team's total wins. Under the point shares system a team's points ($Points = 2 \times Regulation\ Wins + Overtime\ Wins + Overtime\ Losses$) are accredited to individual players based on their contribution to team success. Further, a player's point shares can be split into defensive point shares (DPS), offensive point shares (OPS), and goalie point shares (GPS).

Defensive point shares are calculated by assigning a portion of a team's goals against adjusted for opponent strength to each player based on time on ice, plus-minus, and a position adjustment. Offensive point shares are calculated similarly but use goals created rather than goals allowed. Goalie point shares are assigned using opponent-adjusted goals against with an adjustment for shots against. More importantly though, each numerically represents a player's contribution to his team's success conveniently distinguishing between offensive and defensive play. It follows then that players could plausibly sustain a negative point share value and also that the sum of each player's point share on a team would estimate the team's true points. As can be seen in Figure 1, sum of player point shares slightly overestimates the number of points a team accrues.

The overestimate is reliably minimal and so point shares passed their first test of validity. Despite their strength as a metric system, point shares maintain their drawbacks. They have been criticized for not accounting for even strength play (5v5) versus special teams play (power play or penalty kill), zone start percentage, or penalties; additionally, they have been criticized for understating the importance of assists in offensive production [19]. In essence, their detractors consider them an oversimplification while their proponents tout their simple single-value approach.

1.2 Data

Hockey Reference [1] tracks point share data for the regular season dating back to 1999 when it could first be calculated due to the introduction of time on ice (TOI) as a recorded statistic to the NHL. However, playoff point share data are conspicuously absent. All data acquired for this analysis were scraped

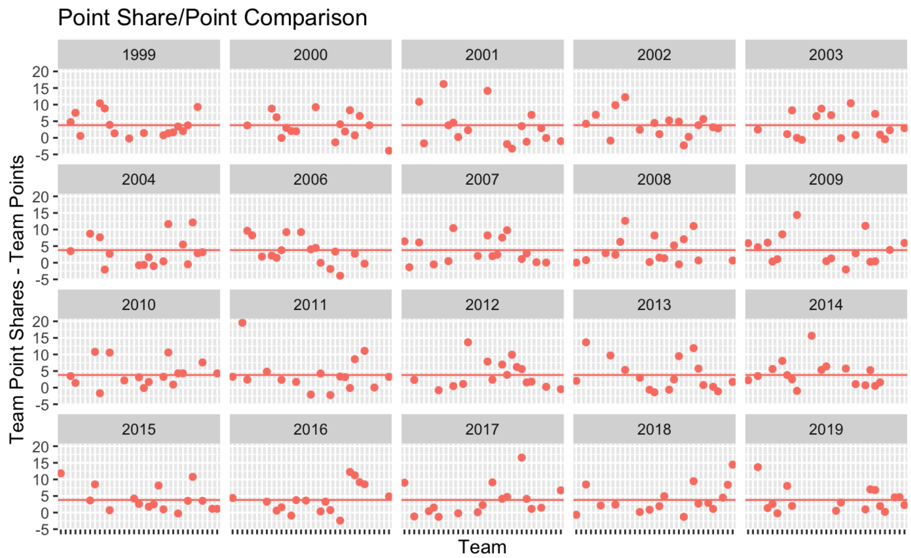

Figure 1: Team playoff points were scraped directly from Hockey Reference while point shares were calculated using methods described later in this paper. Yearly means represented by horizontal bars illustrate that point shares overestimate a team's points by an average of 3.8 points. This is either due to a bias implicit in the point shares metric or a bias in point shares calculation. Either way, the bias is minimal and consistent.

from Hockey Reference [1]. Detailed documentation of the scraping process can be found on GitHub [7]. The result of the scraping process was six tables: three each for playoff and regular season data. A 'teams' table housed team-aggregated data (games, wins, overtime wins, overtime losses, goals, goals against). A 'skaters' table housed individual skater playoff data (position, goals, assists, plus-minus, time on ice). Finally, a 'goalies' table housed individual playoff goalie data (goals, shots, minutes). The process was then repeated for regular season data. All entries were labelled by year and team for easy indexing. After removing rows with missing values and using a join to encode team names as a three letter identifier, the playoff and regular season data sets contained 320 and 597 observations of teams, 6,797 and 7,763 observations of skaters, and 497 and 1,898 observations of goalies, respectively.

2 Methods

At its most basic, this analysis was aimed at partitioning Stanley Cup playoff performances into defensive and offensive components and measuring the relative importance of each to playoff success. Point shares provided a medium through which to classify and quantify player performance. However, since playoff point shares were not recorded in Hockey Reference, they had to be calculated before the modeling process could be implemented.

2.1 Point Share Calculations

Each of the scraped playoff data tables were used in the calculation of point shares [7]. Justin Kubatko's step-by-step calculation instructions [9] were meticulously implemented, resulting in each individual skater having a score for DPS, OPS, and GPS unique to each season in which they played. For non-goalies, GPS is always zero and vice versa. The distribution of the deviations between calculated values and scraped values can be seen in Figure 2.

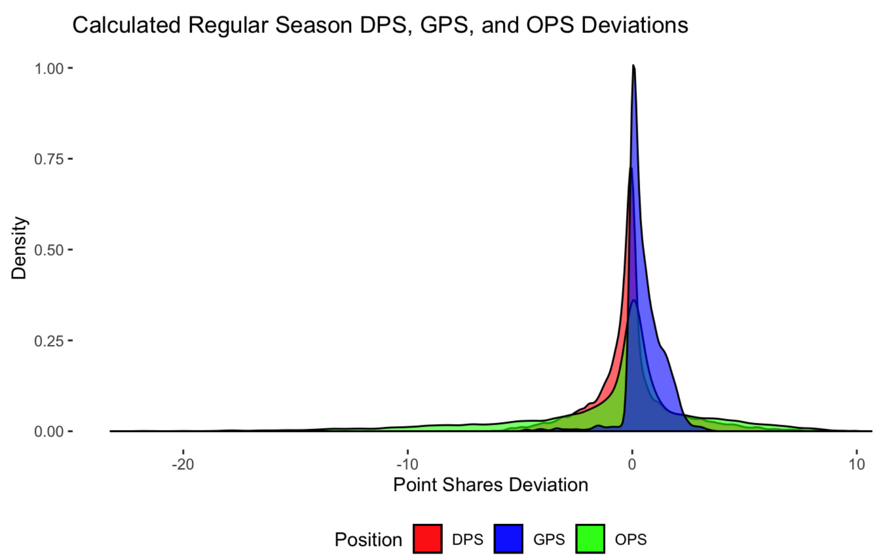

Figure 2: To verify calculations, the calculated regular season point shares were subtracted from scraped regular season point shares. While the distribution of deviations is certainly clustered around zero and serves as a partial validation, there is some concern associated with the variance of OPS deviations as they are relatively large, especially in the negative direction. This is likely due to an uncorrected bias in the calculations script and likely explains the bias in point shares as an estimator of team points.

Then, point shares were aggregated across a team and a season yielding team DPS, OPS, and GPS scores. Originally, these scores were to be turned into three proportions representing the fraction of a team's total playoff points attributable to each position. This could then be adjusted relative to seasonal averages yielding a metric that describes a team's defensive, offensive, or goalie-mindedness compared to that year's playoff field. Unfortunately, by virtue of point shares being possibly negative, proportions lose meaning quickly and seasonal adjustments brought most scores too near to zero to be differentiable. Raw scores could not be used either as teams that progressed further into the playoffs had more opportunities to accrue points, and therefore, point shares. Thus, scores are adjusted relative to a team's number of games played in the playoffs, yielding average DPS, OPS, and GPS scores per game for each team in each season.

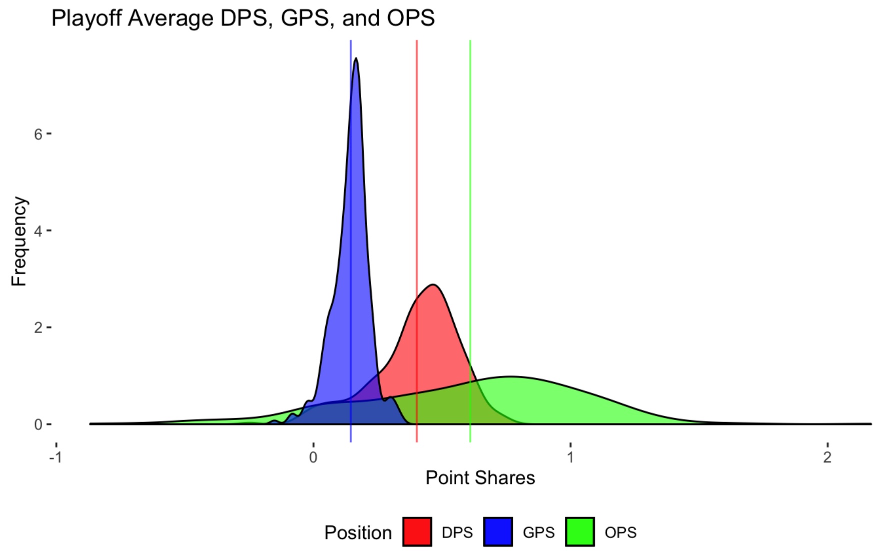

Figure 3: The distribution of calculated playoff team average point share values per game. Interpreted relative to each other, the positional shares explain how team success is typically distributed amongst defensive, offensive, and goalie components. It can be seen that DPS and GPS are distributed with little variability compared to OPS which sees the greatest and also most negative values. However, it can also be seen from the vertical bars representing group means that offensive play, on average, dominates both defensive and goalie play.

Individually, the scores lack much interpretability; however, when considered as a group, the scores give the average contribution of each position to a team's success relative to the other positions.

2.2 Modeling

Average point share values were then fit to a logistic regression model against an indicator variable representing whether a team won the Stanley Cup in their given season. A series of likelihood ratio tests from the lmtest package [23] were used for feature selection [13]. It was found that neither average DPS nor OPS could be removed from the model without detriment. Additionally, average GPS was selected to be combined with average DPS. It makes intuitive sense that

GPS should be included as defensive play rather than offensive play, a sentiment shared by the ratio test. Therefore, average DPS and GPS were added for the final model. Let $\pi_i = \mathbf{E}(Champ_i|(DPS_i + GPS_i), OPS_i)$. Then,

$$logit(\pi_i) = log(\frac{\pi_i}{1-\pi_i}) = \eta = \beta_0 + \beta_1(DPS_i + GPS_i) + \beta_2 OPS_i \tag{1}$$

3 Results

The logistic regression model was fit using the glm() [17] function in R [18]. β_1 and β_2 were estimated to be 11.5 and 6.9, respectively, with p-values less than 0.001.

3.1 Model Adequacy

Aside from significance, the model was primarily assessed using a Hosmer-Lemeshow Test performed using the ResourceSelection package [10] in R. The Hosmer-Lemeshow test partitions data into groups and compares expected and observed successes within each group in what amounts to a glorified Pearson chi-squared goodness-of-fit test [13]. With a p-value of .98 using ten groupings, there was little evidence to suggest that the model was not a good fit. Additionally, a psuedo-R^2 metric named McFadden's R-squared was calculated to be .45 using the pscl package [24] in R. Finally, a power analysis was performed using the WebPower package [25] with a method proposed by Dr. Eugene Demidenko of Dartmouth University [2]. The model was unable achieve a power greater than .051, even for high-sample size iterations. Although, a study published by the National Academy of Sciences [11] points out that significant predictors cannot always find strong predictive power and the foremost goal of this analysis was never prediction.

3.2 Model Interpretation

Following conventional logistic regression parameter interpretation, e^{β_1} and e^{β_2} were estimated, each representing the odds ratio or the multiplicative difference in odds associated with a one unit difference in average positional point shares per game. The odds ratio for combined average DPS and GPS was found to be 98,402.2 and for average OPS, it was found to be 962.3. This was a monumental discrepancy investigated by examining the underlying distributions of the variables under scrutiny which can be found in Figure 4.

Calculated OPS are distributed with more variance and at often greater magnitudes than DPS. It is more typical to see a team that is perceptibly better offensively than its opponent than one that is perceptibly better defensively. Therefore, a one unit difference in average DPS would yield a much higher odds ratio (98,402.2) than a one unit difference in average OPS (962.3). For a more statistics-based interpretation that accounts for difference in standard errors, the

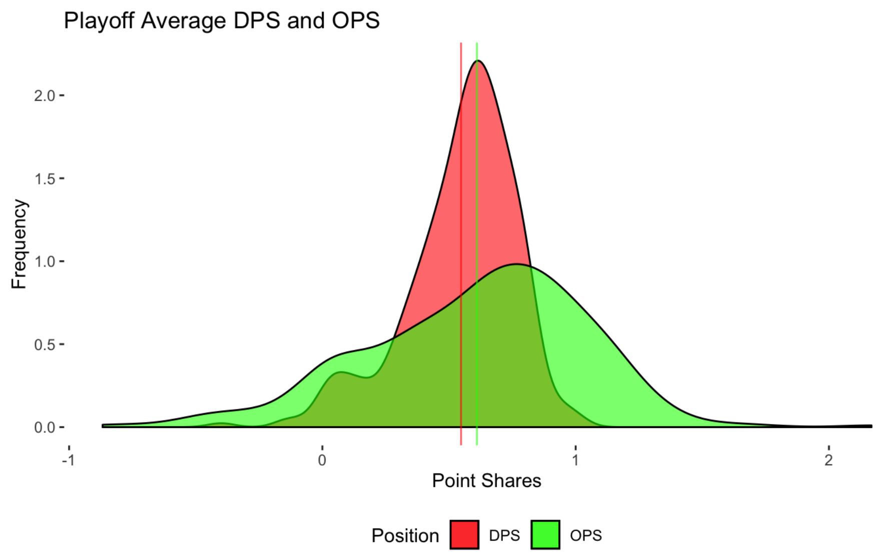

Figure 4: The distribution of calculated playoff team average point share values per game now combining DPS and GPS to match the structure of the fitted model. Offensive play still, on average, dominates defensive play but the difference is reduced since the inclusion of GPS into DPS. OPS is also still distributed with higher variance, leading to a smaller odds ratio for OPS than DPS.

odds ratios for a one unit difference in Z scores were calculated as well and found to be 64.7 for DPS and 106.1 for OPS. While this is certainly an enticing metric, it loses credibility as its real-world interpretability is diminished. Instead, the difference in average DPS and OPS required to double a team's odds of winning the Stanley Cup was derived by first calculating the log-odds ratio between an x difference in the predictor and an x' difference in the predictor:

$$\begin{aligned}
log(OR') &= \eta(x+x') - \eta(x) \\
log(OR') &= [\beta_0 + \beta_1(x+x')] - [\beta_0 + \beta_1 x] \\
log(OR') &= [\beta_0 + \beta_1 x + \beta_1 x'] - [\beta_0 + \beta_1 x] \\
log(OR') &= \beta_1 x'
\end{aligned}$$

Solving for β_1 yields:

$$\beta_1 = \frac{log(OR')}{x'}$$

The log-odds ratio for any x'' difference can be given by:

$$log(OR'') = \beta_1 x''$$

Substituting β_1 into this log-odds ratio gives:

$$log(OR'') = \frac{x'' log(OR')}{x'}$$

Solving for x'' yields an equation that achieves the above desired result by calculating the required difference in a predictor to double a team's odds of winning the Stanley Cup if 2 is plugged in for OR'':

$$x'' = \frac{x' log(OR'')}{log(OR')}$$

When applied to average DPS and OPS, it was found that an average 0.06 point difference in DPS per game doubled a team's chances of winning the Stanley Cup and a 0.1 point difference in average OPS per game had the same effect.

3.3 Results Interpretation

In context of answering the question of whether or not exemplary defensive play wins championships in the NHL, interpretation becomes even more challenging. The original odds ratios show that a team's odds of winning the Stanley Cup can be astronomical for superior-enough defensive teams. However, they also suggest that this occurrence is a rarity and that, relative to offenses, defenses and rather uniform. The Z-score adjusted odds ratios might imply that offensive play is more valuable than defensive play, however the interpretation is extremely muddled. The final metrics can be interpreted again as evidence that a small defensive advantage can be as powerful as a large offensive one. It must be remembered, though, that significant defensive advantages are rare and that typically, outstanding offensive performances dominate defensive play, likely due to the increased pace and aggression of playoff hockey.

It is easily disputed whether defenses or offenses win championships. What is more difficult to dispute, though, is that both win championships. Figure 5 visualizes in the most stripped-down language possible a team's defensive performance, offensive performance, and playoff success. With seemingly no regard for temporal trends, teams that are skilled defensively and offensively tend to find playoff success. Teams that are balanced without being exceptional find more success that exceptionally unbalanced teams. Winning a Stanley Cup, it seems, is a product of defensive and offensive strength with a slight weight towards offense and a strong advantage available to the rare defensively-extraordinary team.

4 Discussion

While carefully designed, this project was partially hampered by metric inadequacies and script inconsistencies. One of the first questions posed during the

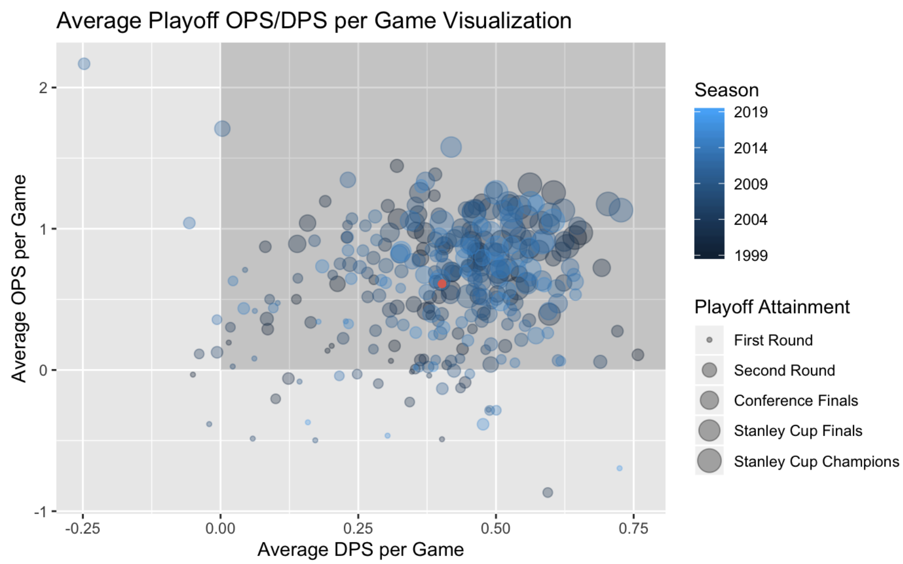

Figure 5: Average DPS plotted against average OPS with temporal and success data represented as visual features of points. The mean is shown in pink and GPS scores are included in DPS. Balanced teams with strong average DPS and OPS tend to find playoff success as can be seen through the positive-sloping trend of playoff attainment. Unbalanced teams (even those that are exceptional in one category) are typically unsuccessful. This visualization implies a parity between the importance of defense and offense and also stressed the importance of both to a championship team.

conception of this project was how to differentiate between defensive and offensive play in a sport where the transition between the two is so continual and fluid. The point shares system was proposed [9] as a potential solution for quantitatively separating defensive, offensive, and goalie play for each individual skater but was also criticized [19] for being an oversimplification, for underestimating the impact of assists, and for placing an undue weight on plus-minus as a defensive statistic. Additionally, due to data-keeping restrictions, point shares could only be calculated as far back as 1999. An evaluation of point shares as an estimator of team points found that point shares consistently but slightly underestimate points; an evaluation of calculated point shares as a reflection of those found on Hockey Reference was performed by comparing the scraped regular season point shares to calculated regular season point shares. While calculated DPS and GPS were accurately distributed, calculated OPS was distributed with greater variance especially in the negative direction. A similar phenomenon can be seen in Figure 6 where playoff point shares are plotted against regular season point shares. The results of this paper must be taken with the assumption that the underlying calculations of OPS may be slightly biased or misguided.

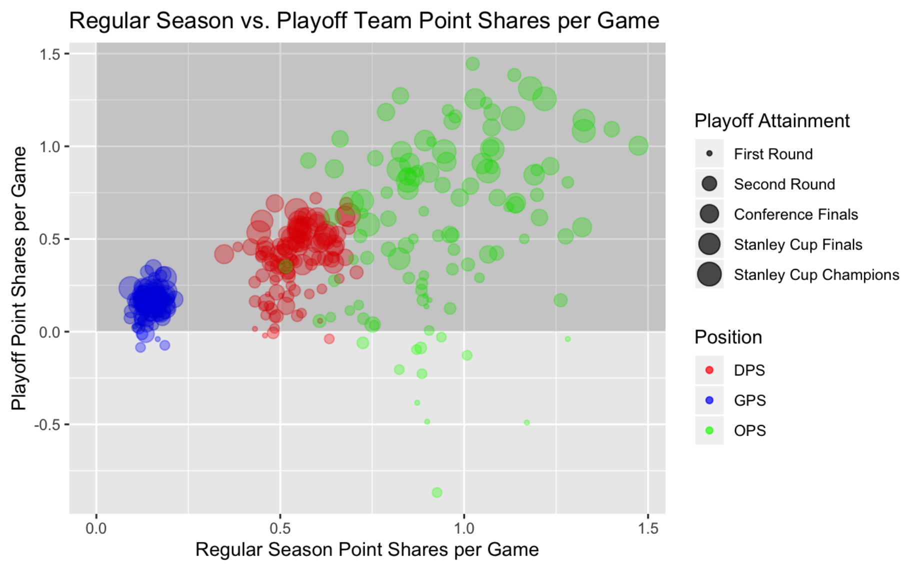

Figure 6: Playoff average DPS, OPS, and GPS per game plotted against regular season average DPS, OPS, and GPS per game. The tight clustering of DPS and GPS values indicates both consistency between regular season play and playoff play as well as consistency in point share calculations. The loose variance of OPS values illustrates the aforementioned bias in OPS calculations. However, it could also be a reflection of a difference in offensive importance between the regular season and the playoffs.

Despite introducing a novel approach to the question of defenses and championships, this analysis fell short of a satisfactory answer as many others did before it. Following the trajectory of this paper though, focus should next be applied to a statistic named Goals Versus Threshold (GVT). GVT was devised by Tom Awad of Hockey Prospectus and operates under a similar system to point shares. Rather than using points, however, GVT uses goals and also a more intricate system of accounting and a more interactive effect between defensive players and goalies [21]. Much of the same analysis featured in this paper could be applied to GVT as it can be broken down into positional components similarly to point shares. It was the hope that GVT could be included in this paper alongside point shares, but time constraints prevented that from being realized. Still, GVT should be considered in all future explorations of the question of whether defenses really do win championships in the NHL.

Acknowledgements

This paper was written under the mentorship and assistance of Dr. Bailey Fosdick [3] and Connor Gibbs [4], and funded by Colorado State University.

References

[1] Hockey reference.

[2] Eugene Demidenko. Sample size determination for logistic regression revisited. Aug 2007.

[3] Bailey Fosdick. Bailey fosdick. 2020.

[4] Connor Gibbs. nfl-draft-trades. 2020.

[5] Justin Glock. Stanley cup playoff hockey vs. regular season hockey. Jun 2012.

[6] Brian Goff. Defense wins championships in the nfl: Fact or folklore? Jan 2019.

[7] Adam Kiehl. nhl_defense. 2020.

[8] Karen Kos and John Albinson. The divergence in perceptions of acceptability of physical behaviors among professional hockey players. *ProQuest Dissertations Publishing*, Jan 2002.

[9] Justin Kubatko. Calculating point shares.

[10] Subhash R. Lele, Jonah L. Keim, and Peter Solymos. *ResourceSelection: Resource Selection (Probability) Functions for Use-Availability Data*, 2019. R package version 0.3-5.

[11] Adeline Lo, Herman Chernoff, Tian Zheng, and Shaw-Hwa Lo. Why significant predictors aren't automatically good predictors. Nov 2015.

[12] milkplantation. Does offense or defense win championships: A statistical analysis. Oct 2018.

[13] Douglas C Montgomery, Elizabeth A Peck, and G Geoffrey Vining. *Introduction to Linear Regression Analysis*. John Wiley Sons Inc., 2012.

[14] Tobias J Moskowitz and L Jon Wertheim. Does defense really win championships? Jan 2012.

[15] Mark Otten. Does defense actually win championships? Jan 2018.

[16] Ijay Palansky. In the nhl, defense doesn't actually win championships. May 2014.

[17] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2019.

[18] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2019.

[19] reckoning. Point shares? (warning: Thread for stats dorks :)). Mar 2011.

[20] J.J. Regan. Why are the stanley cup playoffs so much different than the regular season? Apr 2019.

[21] CJ Turtoro. Catch-all statistics part i: Gvt versus point shares. Aug 2014.

[22] Travis Yost. Offence wins championships in today's nhl. Apr 2019.

[23] Achim Zeileis and Torsten Hothorn. Diagnostic checking in regression relationships. *R News*, 2(3):7–10, 2002.

[24] Achim Zeileis, Christian Kleiber, and Simon Jackman. Regression models for count data in R. *Journal of Statistical Software*, 27(8), 2008.

[25] Zhiyong Zhang and Yujiao Mai. *WebPower: Basic and Advanced Statistical Power Analysis*, 2018. R package version 0.5.2.