Linus Ullmark, Jusse Saros, or Ilya Sorokin: Evaluating NHL Vezina Trophy Voting

Matthew Sheridan

May 6, 2023

1 Introduction and Literature Review

In short, hockey goalies are weird. Who willingly wants to put giant pads on and stand in front of 90 mph shots by players who train every single day to shoot harder and more accurate shots? Despite this, there is a small subset of individuals who do it very well, and are handsomely paid to do so. The best of these in the National Hockey League wins the Vezina Trophy, and award that "is presented annually 'to the goaltender adjudged to be the best at his position'" as voted on by the NHL General Managers (*NHL Awards: Vezina Trophy*, N.d.). Although there is surprisingly little literature on what the GMs use as criteria for voting, it's the worst kept secret that simple statistics such as goals against average, wins[1], save percentage, and games played[2](Woodling, 2019). This analysis is certainly necessary, as the general managers are typically a group who are somewhat adverse to the social pressures of award voting, yet are members of an 'old boy's club'[3] and their perception of the game is somewhat dated.

The most comprehensive statistic that the hockey analytics community uses to evaluate goaltenders is Goals Saved Above Expected (GSAX). GSAX is a statistic that simply takes the expected goals against a goaltender and subtracts the actual goals allowed, which is a measure of their performance (*Moneypuck Glossary*, N.d.). To get expected goals, we will be using Moneypuck.com's[4] model's results, as I do not have the computing power to create my own model. According to Moneypuck, "this model predicts the probability of each shot being a goal. Factors such as the distance

[1]Here, I'm using wins per game played because there were two seasons shortened due to a lockout and then COVID-19, which changed the number of games from the traditional 82

[2]As before, I'm using games played per season length due to shortened seasons

[3]An old boy's club has many meaning, but in sporting terms it's traditionally a group of people (usually men) whose employment perpetuates itself - that is to say the same group of 50ish people are generally the only ones in consideration for these jobs and it's difficult to break into the sphere.

[4]Moneypuck.com is a popular hockey analytics website. I could have also used Evolving-Hockey's model, but Moneypuck had data that was more accessible.

from the net, angle of the shot, type of shot, and what happened before the shot are key factors in the model" (*MoneyPuck Shot Model*, N.d.).

Unfortunately, the NHL does not publicly release their puck tracking data, unlike other leagues. All of the teams have access to this data and use it to create their own expected goals models. If we had this data, we would have far more accurate models that would include passing data as well[5].

To create a WAR model, I'll also need to quantify exactly what contribution a goaltender has to a win. To do this, I'll use Evolving Hockey's Goals Per Win (GPW) metric, which uses a weighted average[6] of the coefficient regressing goal differential per game (dependent) on wins per game (independent) (*Evolving Hockey Goals above replacement*, N.d.).

I'll be combining Goals Saved Above Expected with Goals Per Win, along with a novel adjustment that relates goaltender performance to team strength. This metric will be the percentage of wins that a goaltender has as a function of his team's wins (wins per team win or WPW). This statistic is of my own creation, and downweights the initial WAR value determined for each goaltender to reflect the amount of actual wins he contributed to his team wins.

I believe in this metric because of intuition[7]: If a goaltender has an amazing statistical season, but is on a really good team, then their backup goaltender will likely also have a high number of wins relative to games played and therefore we'd be downweighting the WAR of a goalie on a better defensive team. On the converse, if we have a goaltender with great stats on a bad team, it's likely that a replacement level goaltender (his backup) will not get many wins because the team in front of him is bad, so the better goalie will be downweighted less because his being on the team contributed more to a team winning than the backup. This essentially helps goalies who play a lot of games on bad teams, and punishes those who win a lot of games on an already good team.

In this paper, I'll be exploring to what extend the classic goaltender statistics influence and can predict voting, trying to predict which goaltender will win the Vezina trophy in the 2022-2023 season. Then, I'll be using team level data combined with Moneypuck.com's GSAX model to create my own WAR statistic, whose relative rankings along with my industry knowledge will determine who should have won the Vezina in past seasons, and who I think should win for the present season (*MoneyPuck Goaltender Statistics*, N.d.). I hypothesize that some combination of Ilya Sorokin, Jusse Saros, and Linus Ullmark will show up as the top goalies for this year. I also believe that Linus Ullmark will be the predicted Vezina winner based on our regression of stats and past sesasons (using the typical

[5]For example, if you make a pass across the face of the goal and get a shot off, it's a lot more likely that you'll score than just simply shooting from that spot. Current models can only base goal probabilities on shooting location.

[6]This weighted average weights the three most recent seasons at 3, 2, and 1 (most recent the highest).

[7]We can see this in practice in the 2022-2023 season - Linus Ullmark has over 40 wins and the best stats, but on one of the best rosters of all time where his backup also has a ridiculous record and stats. Another goalie, Jusse Saros, has insane stats and most of the wins on a poor defensive team, and therefore should not be punished because his team doesn't win a lot.

GM focused stats). Finally, I conject that when discounting using team strength and my WAR model, Ullmark will not be the predicted Vezina winner, and that either Connor Hellebuyck or Ilya Sorokin will be the predicted winner.

2 Data

To work with this data, I used Excel to scrape websites for data, manually manipulated data in Excel, and used the following R packages to help manipulate the data, analyze the data, and create models: broom, dplyr, ggplot2, MASS, stargazer, tidyverse, and xtable.

To help analyze the prior Vezina voting landscape, I gathered data for many different goaltender statistics from 2008-09 to 2022-23 using Moneypuck's data page (including the lockout shortened 2012-13 season and the COVID shortened 2019-20 season) (*MoneyPuck Goaltender Statistics*, N.d.)[8]. I combined this with Hockey Reference's Vezina Trophy voting counts to get the dependent variable (*Hockey Reference Vezina Trophy*, N.d.)[9]. I merged this data with Quanthockey's goaltender statistics for the top 50 goalies in games played from the years of 2008-09, mainly wins and games played (*NHL Goalie Stats Data*, N.d.)[10]. Since I was using win percentage in games played and games played percentage, I was not able to use any goalie who was not in the top 50 for games played as I lacked the knowledge and ability to be able to web scrape the multiple pages of QuantHockey's data efficiently. Thankfully, not including this data did not change the fundamental distribution of goals against average and save percentage but removed some outliers that would have been extremely noisy in the dataset, as seen in figure 1.

Our next set of data combined two sources as well. This was to create the goals per win statistic. First, I took a table of data from Evolving Hockey with already calculated GPWs for 2007 until 2020 (*Evolving Hockey Goals above replacement*, N.d.). This was missing both this year and last, so I had to manually calculate the last two values. To do this, I used data from QuantHockey on season standings starting at 2016 that included a goal difference per win column, and were able to use it to run the same weighted regression and generate two goal per win values for 2021-22 and 2022-23 (*NHL Standings Data*, N.d.)[11]. After doing this, I also generated a couple extra values for the years before 2021 and 2022 to see how accurate it was, and although the values that were compared weren't the same, they were all very close so I opted to use a combination of Evolving Hockey's goals per win table and our two imputed values, shown in table 1.

Figures 2 through 7, in the appendix, show the six boxplots summarizing the data at hand (Goals

[8]Data from 2008-2022 available here
[9]All voting data accessible from this page
[10]All year goalie stats accessible from this page
[11]All season standings available from this page

Against Average, Save Percentage, Games Played Percentage, Wins per game played, and Vezina Votes). Table 2 shows the summary statistics of the variables of interest. We can see that the votes are highly right skewed, and due to this variance we will not be able to use a linear model to model this. Among vote receivers, votes are still right skewed but have a somewhat more interpretable distribution Furthermore, games played percentage, save percentage, goals against average, and wins per game played are all pretty normal without any largely influential outliers.

This analysis also used the top 50 goalies in games played per year because those were the ones we were able to get wins for, and the analysis is limited to the years 2016 to 2022, because that's all we're able to get standings data from.

3 Vezina Trophy Voting Analysis and Predictions

3.1 Methodology

The regression I used was negative binomial regression, which I used because the Votes category is a count variable and has a very high variance. Our data was all goalies from years that were not 2022, as the voting has not commenced for 2022.

The formula was: $log(Votes) = \beta_0 + \beta_1 GAA + \beta_2 SVPCT + \beta_3 WGP + \beta_4 GPPCT + \epsilon$

Where the β's are the intercept and coefficients and the ϵ is the unobserved error term.

3.2 Results

The regression output is shown in table 3.

The coefficients are the contribution of a unit change in each variable to the log of the predicted votes for a goalie. We can see that goals against average is insignificant even at the 10% significance level, but the coefficients of wins per games played, games played percentage, and goals against average are all statistically significant. Although it's hard to compare the coefficients of GAA and the others, it seems that save percentage has the highest effect, although we know it is centered around 0.91 and has a low spread, while games played percentage and wins per game percentage all have a mean around 0.5 and a large spread.

After creating this, I used the model to predict the number of votes on the 2022 crop of goalies, all who played in the top 50 of games played. To make this interpretable, I linearly scaled all the predictions by a factor of $\frac{288}{\sum predictions}$. This is because there are 288 votes (33 GMs with a 5-3-1 voting system), so we can scale the predictions to the correct voting scale (*NHL Awards: Vezina Trophy*, N.d.).

Figure 8 presents the top 10 goaltenders in predicted votes for the 2022-23 season, while figures 9 and 10 show the top 10 scaled predictions for 2 other years (2013-14 and 2018-19).

These predictions, unsurprisingly, have Linus Ullmark, who is the best statistical goaltender in the salary cap era (but played behind the best defense ever) as the lights out favorite to win the trophy, which is to be expected. Unfortunately, there is not a plain and easy way to control for the maximum number of votes a player can receive is 165 and they all sum to 288, but we'll work through the limitations. Furthermore, I also note that for every year except the 2013-14, 2018-19, and 2019-20 seasons, the predicted highest vote getter was also the winner.

4 Wins Above Replacement Model Description and Results

4.1 Method

My new wins above replacement statistic will be created with the following formula:

$$(GSAX/GPW) * WPW$$

where GSAX is goals saved above expected, GPW is goals per win, and WPW is the goalie's wins as a percentage of his team's total wins.

To do this, I took the Moneypuck GSAX data and, for each year, divided every goalie's GSAX by the GPW, then multiplied by each individual goalie's WPW (*MoneyPuck Goaltender Statistics*, N.d.). The following table shows the distribution of WAR values.

Year	2016	2017	2018	2019	2020	2021	2022
Min	-2.325	-3.711	-3.295	-2.211	-2.421	-4.283	-1.922
Max	5.100	4.247	1.910	3.613	2.817	4.496	6.619
Median	0.15689	-0.03994	-0.11480	-0.15629	-0.36039	0.03603	0.14224
Mean	0.47830	0.11106	-0.07856	-0.30621	-0.14597	0.35577	0.72105

4.2 Results

We can see that the WAR values are all mostly centered a bit above zero, which makes a lot of sense because we used the top 50 goalies in games played, who are likely better than other goalies, and it involved likely at least one goalie on each team. Furthermore, The minimum values range from about -4 to -2 and the max values range between 2 to 7. As a sanity check, hockey teams usually win between 25 and 55 games, and a goalie who plays about 50 games adding about 3 wins of value is not absurdly low or high, as there's another goalie and 20 skaters on the roster.

In every year except for 2017, the leader in WAR won the Vezina trophy (*Hockey Reference Vezina Trophy*, N.d.). It is tough to explain what happened in 2017, but the best I can do is that the goaltender, Antti Raanta, who was the leader in WAR, was on the Coyotes who are a small market, underperforming team that won 29 games (Raanta had 21 of those wins) and he basically kept the Coyotes in most games and they gave up a lot of high danger chances.

Interestingly, however, we're concerned with how our WAR model relates to this year's crop of goaltenders. The table below shows the top 5 rankings of goaltender WAR in 2022.

Name	Goalie WARs
Juuse Saros	6.619963
Ilya Sorokin	5.220237
Linus Ullmark	4.730871
Igor Shesterkin	3.983068
Connor Hellebuyck	3.966646

As the title of this paper suggests, we're interested in Saros, Sorokin, and Ullmark[12]. Saros had a huge workload and dragged his team to the second last game of the season in playoff contention. Sorokin also dragged a subpar team to the playoffs, where they lost in 6 games. Ullmark was the best goalie in terms of GSAX and every other classic stat, but he is actually punished because his team was so good and had 65 wins, while he only had 40 of them.

My WAR model takes GSAX, one of the best goalie stats out there, as well as a weighting of what percentage of wins a goalie contributes to his team, and assigns them a value of wins added in a season.

Thus, contrary to the previous section, Ullmark probably should not run away with the Vezina Trophy in the 2022 season. The WAR model defintely corrects for the worload a goalie faces: If you're on a bad team you should be rewarded a lot more for performing really well and your performance should be discounted if you're performing well on a really good team.

5 Discussion

Evaluating NHL goaltenders is pretty similar to evaluating NFL running backs. They're a position whose stats are heavily dependent on the team in front of them, and it's hard to isolate their impact. Take the 2022-23 Boston Bruins - a team with a stacked roster that broke the all time record for wins. Yes, they had some great goaltending seasons, but do I really think they had 2 top 10 goaltenders on the year? The answer is probably not. They were an insane defensive team that not only had luck on their side (until the playoffs), but also had some great goaltending.

[12]The title of this paper was created before this model was run. Yes, I watch a lot of hockey

The 2022-23 Vezina Trophy is going to go to Linus Ullmark, mostly because of his baseline statistics (GAA, SVPCT, Wins). I believe that the analysis I did in part 3 is a lot more in line with what actually will happen with the trophy - Linus Ullmark will run away with it (Woodling, 2019).

Although it's a bit hard to describe the influence of each predictor (SVPCT, GAA, GPPCT, WGP) on the predicted votes, we know that controlling for all four variables makes the Vezina winner extremely predictable. Running a regression of Votes on GSAX does not provide meaningful results, as some goalies have high GSAX but are just on horrible teams and have bad basic stats, including wins, which are generally accepted as more of a team statistic by the analytics community.

While this paper is titled "Linus Ullmark, Jusse Saros, or Ilya Sorokin: Evaluating NHL Vezina Trophy Voting," the three finalists haven't been announced yet. They will most likely be those three, and I believe my WAR model quantifies that. For some anecdotal justification - Ullmark had a great season but was behind a great team. I think he should be nominated but not win. Saros carried a subpar, injury riddled Nashville Predators squad to within 2 wins of a playoff spot, playing a heavy amount of games and making otherworldly saves. Ilya Sorokin, who is actually my pick to win the award had I been voting, as he carried a weak, weak Islanders squad to the playoffs whose style involves scoring few goals and trying to suffocate offenses. He was really the only reason they made the playoffs, in my opinion, and exemplifies the description of the award.

Regardless of who wins, the voting process for the Vezina Trophy is dated - a group of GMs who don't like analytics much get into a room and decide which goalie has the most wins, highest save percentage, and largest goals against average. Although turning the vote over to sports writers could shift the narrative in the right direction, I think social pressures could keep the same winners and change very little.

Since I believe that the league's General Managers are probably picking the wrong statistics to value, a more comprehensive statistic and ranking of goaltenders could provide useful, all encompassing information.

The Wins above replacement model I created provides intuitive rankings and follows Moneypuck's WAR rankings to a general degree[13]

Overall, there is more work to be done in this field. Goaltenders are extremely understudied, and it's extremely hard to value them. How is it possible that Boston was able to take a goalie who played few games on a horrible Buffalo Sabres squad, give him a contract him that made him one of the higher paid goalies, and turn him into a Vezina winning goalie? There are so many new avenues to explore with goaltending, and until we have puck tracking data and player tracking data there is only so far the public can take their models.

[13]This was a new feature recently added. Though our rankings are similar, they are not the same and we have some noted discrepancies.

6 Limitations

There were a few limitations that are necessary to address. First off, almost all of our analysis was based on the top 50 goalies in games played per season. This is not ideal, but there was a lot of noise that could have showed up, especially with goalies who play a small number of games due to injury whose stats are really extreme in either direction and emergency backup goalies who played a few minutes of a game and skew more data.

Furthermore, the season by season team data we had was only from 2016-2017 on due to QuantHockey only going back that far, so we were only able to get 5 years worth of WAR projections.

I was also not able to effectively evaluate the WAR model or the Vezina votes model for effectiveness. For the Vezina voting model, the best I could do was evaluate rankings, however the ranking evaluation method is beyond my education and the scope of this paper.[14]

For the WAR model, it seemed like the highest WAR goalies would be high in the Vezina rankings, but didn't always win. There really wasn't a way to explore how 'accurate' it was, as it was a self defined metric that I want to use to evaluate goaltenders. Thankfully, I can say with confidence that the rankings fall in line with the hockey analytic's community's perception of goalies, as well as who will probably be nominated for the award.

In the future, expected goal models will be better and therefore our advanced statistics will be better. I could have also combined GSAX data from other sources (Evolving Hockey, HockeyViz, etc.), however many of the sites providing this are behind a paywall.

I want to thank MoneyPuck, EvolvingHockey, QuantHockey, and the ECON 1042 teaching staff for providing great data and helping me with this project.

[14] But the eye test tells us that our rankings were pretty good...

References

Evolving Hockey Goals above replacement. N.d.

URL: *https://evolving-hockey.com/glossary/goals-above-replacement/*

Hockey Reference Vezina Trophy. N.d.

URL: *https://www.hockey-reference.com/awards/vezina.html*

Moneypuck Glossary. N.d.

URL: *https://moneypuck.com/glossary.htm*

MoneyPuck Goaltender Statistics. N.d.

URL: *https://moneypuck.com/goalies.htm*

MoneyPuck Shot Model. N.d.

URL: *https://moneypuck.com/about.htm#shotModel*

NHL Awards: Vezina Trophy. N.d.

URL: *https://records.nhl.com/awards/trophies/vezina-trophy*

NHL Goalie Stats Data. N.d.

URL: *https://www.quanthockey.com/nhl/seasons/nhl-goalies-stats.html*

NHL Standings Data. N.d.

URL: *https://www.quanthockey.com/nhl/standings/standings.php*

Woodling, CJ. 2019. "It's time to revoke NHL General Managers' voting privileges for the vezina trophy.".

URL: *https://www.anaheimcalling.com/2019/6/20/18692571/time-revoke-nhl-general-managers-voting-privileges-vezina-trophy-gibson-bishop-lehner-vasilevskiy*

Appendix

Figure 1: Comparison of GAA and SVPCT for the dataset with only goalies in the top 50 of GP and the entire dataset

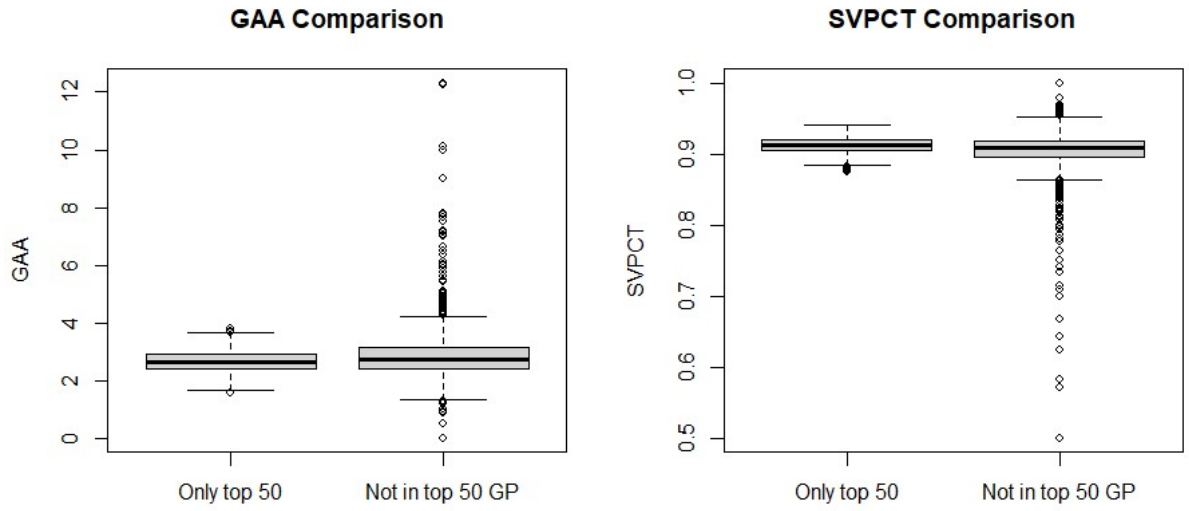

Table 1: Goals per win table. Most data from: (*Evolving Hockey Goals above replacement*, N.d.)

Year	Goals Per Win
2008	5.53
2009	5.53
2010	5.60
2011	5.73
2012	5.39
2013	5.28
2014	5.25
2015	5.18
2016	5.31
2017	5.13
2018	5.36
2019	5.62
2020	5.57
2021	5.25
2022	5.54

Figure 2: GAA distribution

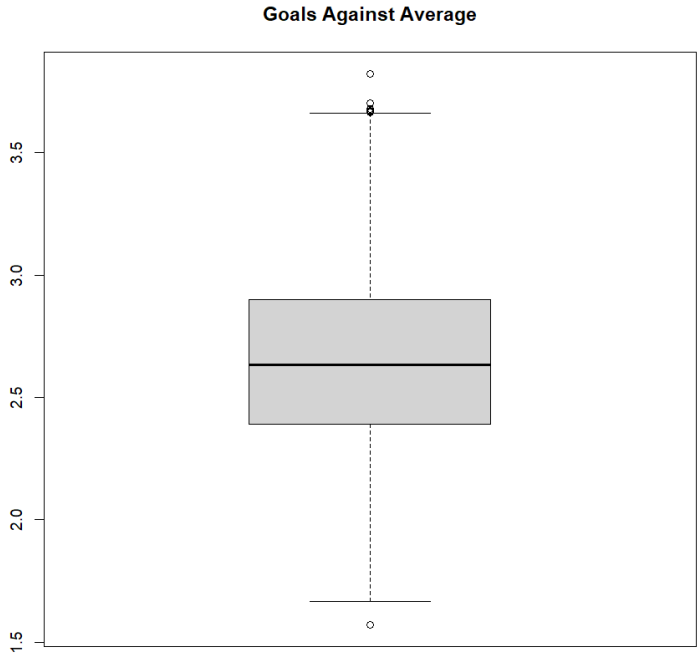

Figure 3: SVPCT distribution

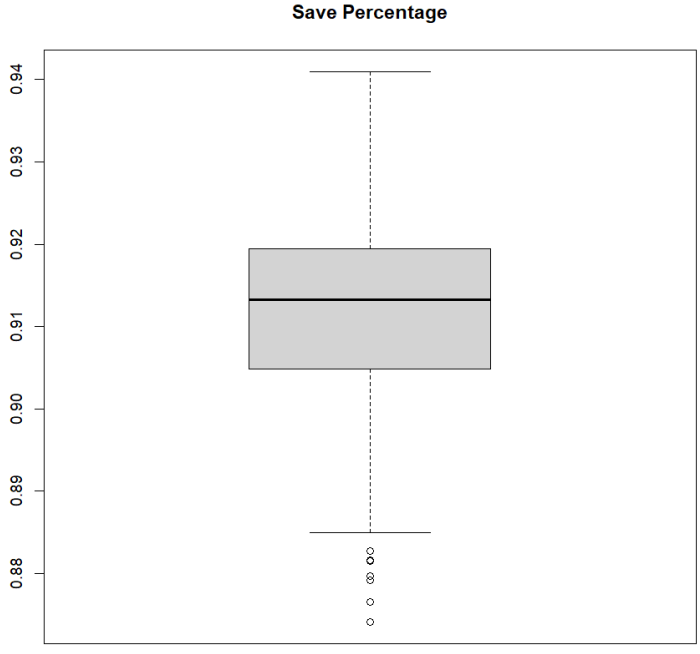

Figure 4: GPPCT distribution

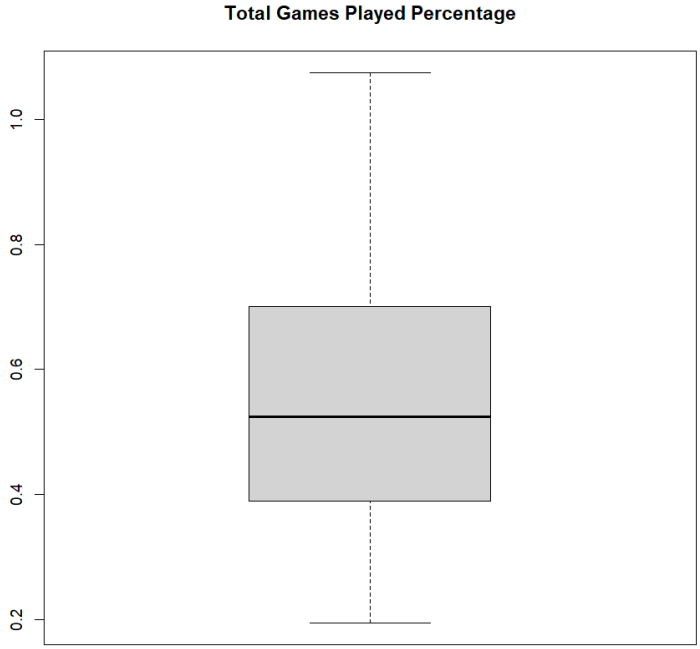

Figure 5: WGP distribution

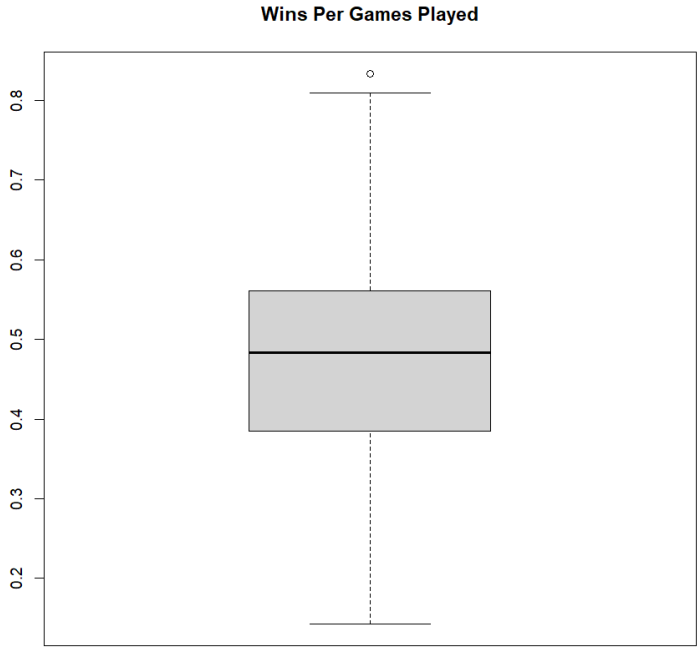

Figure 6: Votes among all goalies distribution

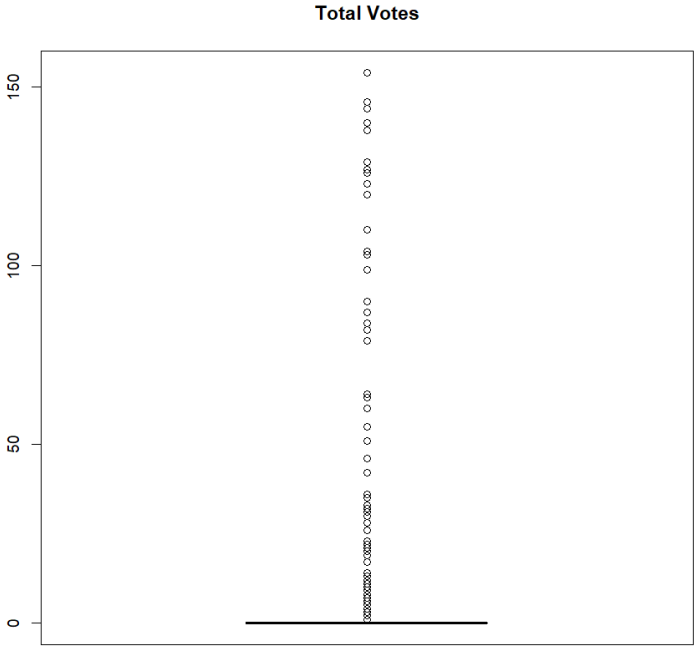

Figure 7: Votes among vote getters distribution

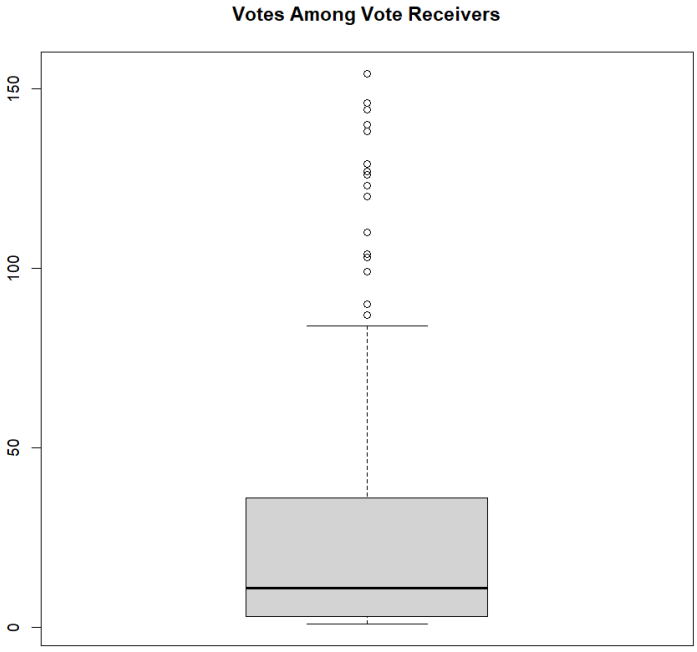

Table 2: Summary Statistics for variables of interest

Votes	GPPCT	SVPCT	GAA	WGP
Min. : 0.000	Min. :0.1951	Min. :0.8741	Min. :1.571	Min. :0.1429
1st Qu.: 0.000	1st Qu.:0.3902	1st Qu.:0.9049	1st Qu.:2.393	1st Qu.:0.3846
Median : 0.000	Median :0.5244	Median :0.9132	Median :2.633	Median :0.4839
Mean : 5.514	Mean :0.5477	Mean :0.9118	Mean :2.658	Mean :0.4741
3rd Qu.: 0.000	3rd Qu.:0.6982	3rd Qu.:0.9195	3rd Qu.:2.900	3rd Qu.:0.5607
Max. :154.000	Max. :1.0738	Max. :0.9409	Max. :3.821	Max. :0.8333

Table 3: Regression Output

	Dependent variable:
	Votes
WGP	16.527^{***}
	(1.797)
GPPCT	8.882^{***}
	(0.814)
SVPCT	219.145^{***}
	(28.396)
GAA	-0.307
	(0.822)
Constant	-215.091^{***}
	(27.792)
Observations	644
Log Likelihood	-596.555
θ	0.239^{***} (0.031)
Akaike Inf. Crit.	1,203.111
Note:	*p<0.1; **p<0.05; ***p<0.01

Figure 8: 2022 Vote Predictions

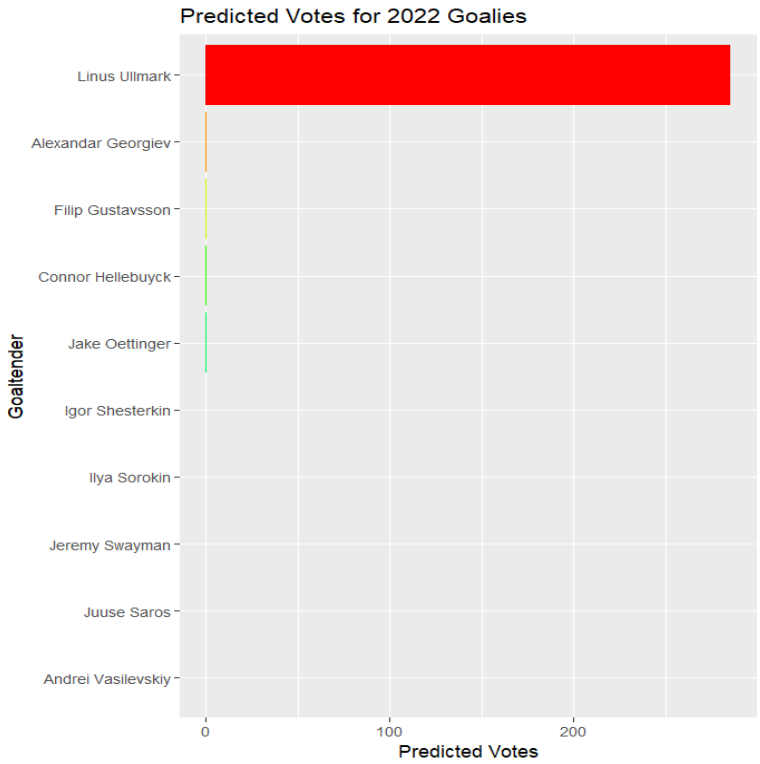

Figure 9: 2013 Vote Predictions (Actual Winner: Tuukka Rask)

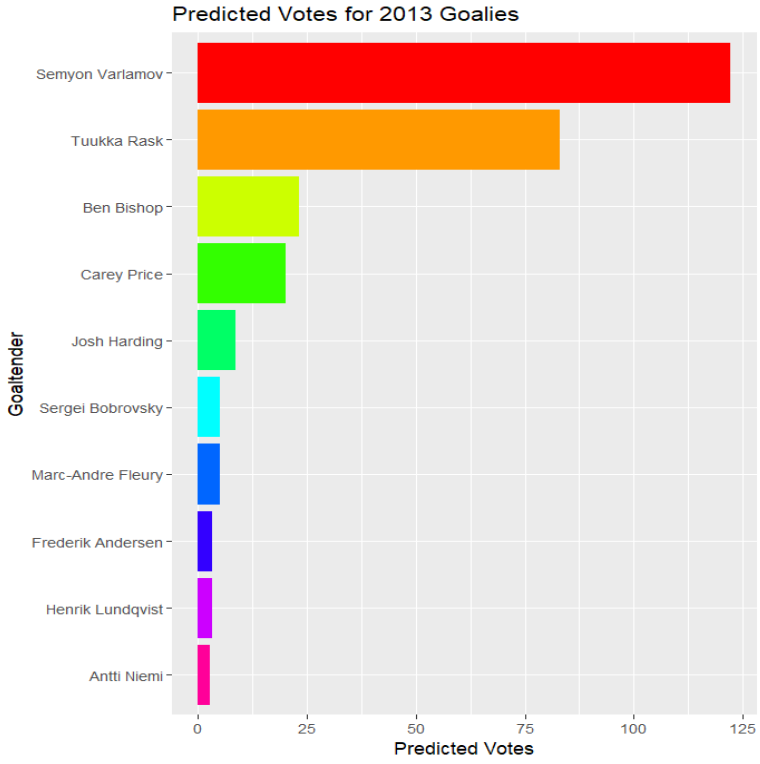

Figure 10: 2013 Vote Predictions (Actual Winner: Andrei Vasilevskiy)