A Naive Analytical Revaluation of Chess Pieces:

Creating a New Way to Learn Chess

Raj Dasani

Department of Data Science, Honors

University of California, Berkeley

May 2022

Abstract

During a chess game, players commonly refer to chess pieces with a certain numerical value: pawns as 1, bishops & knights as 3, queens as 9, etc. These numbers were created from the start of chess, and have not been changed since. These values are used across all players in order to help articulate how well a player is doing during a certain game. But when we compare players at different levels, we see different patterns with how players utilize their pieces and how they value them. This thesis seeks to focus on analyzing how we can numerically quantify chess pieces, and how these valuations can differ across levels.

From creating proportion data to analyzing the odds of logistic regression, this paper will use these techniques to analyze past games in order to learn how different level players use these pieces and where we see either a difference in value or usage. Through these models, we see patterns that grow and fall across different levels, adding to this hypothesis that the current values we use today are catered towards the best masters in the world rather than the average player. Because of this, this thesis will ideally be the start of creating a new way of learning about chess, from a relatively naive perspective of looking at pieces from more of an individual value, opening up the doors to a dynamic way to improve at any level.

Introduction

The game of chess is one of the earliest known games in history. The game has evolved through its different schools throughout the centuries which has led to different ways for individuals to learn and improve at the game. Some of the most popular methods include memorizing openings, completing puzzles, reviewing best moves, and consistent practice (Chess.com). Especially within the last 50 years, these have been paired with the rise of computers and chess engines designed to figure out the next best single move at a certain point in the game. From the famous Deep Blue, an IBM supercomputer, which beat the then world grandmaster Garry Kasporov (History.com Editors) in 1996, to the popular ones available to the public today such as Chess.com's Stockfish (Chess.com), which can give nearly instant feedback to anyone who plays a game on it, people are improving faster than ever as players seek to understand the perfect strategies and patterns. Grandmasters who are able to understand the way these engines play are able to distinguish themselves as the best players in the world because of their ability to replicate groups of moves, along with their understanding of how to incorporate these strategies into their current play. But, with what I dub as a form of "learning from perfection", learning from chess engines can have its faults, especially for younger, inexperienced players, as they can skip steps of understanding the fundamentals of the game (Serper) and simply lean on the engine to teach them the best move. Whether this is a good thing or not, the idea of computers assisting the learning process of chess is an interesting growth in the "chess learning industry", but one that has only been geared towards perfection. What has not really been done, and what I seek to do in this paper, is use data science and machine learning to help us learn about these fundamentals of chess that experts in the field claim is lost, and seek to learn more about the most fundamental part of the game, the pieces and their respective values.

In other words, this paper aims to look at chess from a learning perspective that is not geared towards perfection but rather understanding the value of individual pieces in a variety of rather “naive” ways in order to start building an understanding of the values of these pieces and how they may contest or agree with traditional values understood in the field. But more than that, to gear it towards a learning perspective, this paper seeks to understand the differences between these calculated values for different levels of chess players in order to see if differences appear, implying the underuse/overuse or undervalue/overvalue of them.

This paper will include two main models: the naive “takes and checks” model and the “logodds” model. These models are designed to be as interpretive as possible in order to output proposed chess piece values for the levels of players I have chosen. The naive takes and checks model is made to summarize all the key dynamic moves in a game: when a player is to take an opponent’s piece as well as the number of checks they get on the enemy king. These stats together typically can tell someone the story of the game in a general sense, but ultimately does not explain the other half of the game - the moves that were not done or could have been done. The logodds model takes a different approach and seeks to view the value of pieces just based on the end of game situation. It uses logistic regression to analyze which pieces are most valuable based on the winning percentage advantage a player has if they have one more of such piece compared to their opponent. Therefore, the logistic regression will provide these logodds based on the linear combination of the remaining pieces as predictors using the 0/1 binary win/loss metric. Both of these models will provide us valuable insights, as detailed below.

Overall, these models will be used to articulate and grow our understanding of the value of chess pieces at different levels, and challenge the traditional evaluations of chess pieces, especially the rook and the queen, that have been used for so many decades.

Literature Review

Understanding the material value of pieces in chess is one of the first things taught to every chess player, as it can help give a basic understanding to a beginner of what they should prioritize when attacking and defending in their own games.

The value of chess pieces has had an accepted standard value for many years, values that are considered arbitrary at best in the modern game, and are used as a very general way of understanding the current balance of the game, to a fault. Any generalized evaluation of pieces is always given the caveat of "it depends", as pieces can have a vastly different value depending on its current placement relative to one's other pieces, the opponent's pieces, the point in the game, the abilities of each of the players, etc. The most traditional values of pieces are seen below, and have been widely accepted since about the 18th century when it was made popular by the Modenese School, but there are many suspicions that it may go farther back than that. These values are the following:

Queen - 9

Rook - 5

Bishop - 3

Knight - 3

Pawn - 1

These values are used in traditional chess lingo and communication today, most popularly noted on the side of Chess.com's UI during a game for the ease of player tracking. The only piece not mentioned in the list above is the king, as the king itself is regarded to have an infinite value because it can never be taken, as if it ever would be in a position too, the game would be considered over as it would be "Checkmate."

One of the earliest published versions of a proposed reevaluation is by Howard Staunton, an English chess master who was regarded as the world's best player in the mid 1800s. In his book, *The Chess-Player's Handbook*, he details the following relative value of chess pieces (Staunton):

Queen - 9.94

Rook - 5.48

Bishop - 3.50

Knight - 3.05

Pawn - 1.00

Stanton's analysis is structured on the idea of exchanging pieces, which in other words is when a player "trades" one of his pieces with another by taking his opponent's piece, knowing the piece that he took with will be taken easily in the next move. He uses the pawn here as 1 for reference, with claims such as "The Knight being worth at least three Pawns", giving way to the extra .05 bump in his evaluation. He gives the Bishop and Rook an extra 0.5 compared to the traditional, and the queen near an extra full point for their ability to attack other pieces, and the vision they have across a board.

More recent re-evaluations have been done in many ways, and have involved using computer algorithms and engines, as detailed above. One of the key ways to distinguish the pieces is the time of the game, as popularized by Larry Kaufman, an American Chess Grandmaster who specifically looked at material imbalances of some of the best games in the world for three phases of a chess game: the opening, the middle game, and the endgame. Specifically, his methodology included looking at games with specified material imbalances and noted which specified pieces were present or absent to then find the average difference between performance rating and player rating, and aggregate all similar situations, to isolate a final calculation for “piece value” (Kaufman). He did this across phases and found a general evaluation of the following:

Queen - 9.75

Rook - 5.00

Bishop - 3.25

Knight - 3.25

Pawn - 1.00

Here we see the knight and the bishop still equal, but with a 0.25 bump, in order to establish again, its value over 3 pawns. The rook here is kept constant, but the queen, yet again, is given the most significant bump, at 0.75 than the standard value.

Lastly, and most recently, AlphaZero, a very popular computer program developed by the DeepMind team at Google made headlines when in late 2017 and then again in 2018, with just 4 hours and the basic rules of chess, it played millions of games against itself and then beat the most popular engine, Stockfish, with a record of 28 wins, 72 ties, and 0 losses (Rozman). The DeepMind team then made their own evaluation of chess pieces, via a paper in September 2020, where it aimed to look at alternative rule sets in chess that could maintain balance and provide a rich variety to the game. Within the paper, they used the weights of a linear model to predict the game outcome purely based on the difference in numbers of the pieces. It uses a mean squared error loss function with the expected outcome along with tanh, a hyperbolic tangent function, in order to create an expected value from all 10,000 AlphaZero self-played games (Tomašev et al.). It found the following values for the classical game:

Queen - 9.5

Rook - 5.63

Bishop - 3.33

Knight - 3.05

Pawn - 1.00

One key similarity between all of these re-evaluations, and the valuations done in the past, is that they use the games of the best players, the players who already have a prior understanding of the value of chess pieces and how they should be supposedly valued. The key question inspired by much of this literature then is as follows: will we see similar trends across rating levels? And if not, how do they differ, and what insights could this give us on how differently players play at different rating levels?

Data

The dataset this analysis uses is a collection of 20,000+ Lichess games from a variety of ratings. Lichess is the second most popular chess-playing site in the world, after Chess.com. The dataset I am using was downloaded from Kaggle, and was originally used for the same reason as me, to analyze chess games. It can be found on the GitHub linked to this project under the data folder as “chess_games.csv”. Lichess now releases a massive amount of games per month, all under their open database, within a Creative Commons CC0 license, which will be used for future iterations of this report. Lichess uses their API to collect the games played by their users, and extracts information such as the ratings of each player, script of the game (every move in standard chess notation), and the end result. These data points became the key aspects of my analysis, and will be detailed more here:

Player Ratings

For the purpose of this report, it should be noted that Lichess's rating system is a bit different than Chess.com's and the standard chess rating systems, as they have their own independent rating system, which many experts determine as "inflated" (Vargas). This still allows us to create the levels of players desired to understand patterns across different levels, but these may need to be inflated from the standard rankings in order to match Lichess's ratings. For simplicity sakes, this research does not account for the inflation and keeps the ratings at face value when playing them into categories. So, the ratings categories are the following:

2500+ : Grandmaster	**1800-1999 : Class A**	**1200:1399 : Class D**
2200-2499 : Master	**1600:1799 : Class B**	**900:1200 : Novice**
2000-2199 : Expert	**1400-1599 : Class C**	<900 : Rookie

The one bolded are the ones I focused on for my analysis for a couple reasons. Firstly, as detailed before, I chose to focus on the levels of players, not deemed as "masters" or "experts" in order to see if we can uncover a difference to the traditional ratings with less skilled players. Additionally, the dataset I used was heavily weighted in size to these ratings, such that there were not enough games for the Expert or Master rating categories (<1000 games). The Rookie rating category had enough games, but also has a lot of variance due to it being the lowest category as well as it being before a point where players truly know the value of pieces. Therefore, it is included in some of the analysis, but was omitted from some of the models.

Chess Script

The chess script is the heart of this project, and what sets up the models detailed later in this paper. For every chess game, there is a script which enlists all the moves that happened in the game by white and black in algebraic notation. This notation is widely known among chess players and is standardized in order for any player to be able to look at a game and understand what happened with the corresponding move. However, it has its specific notations that are obvious to understand as a chess player, but can be very tough to understand in plain code. The entire script is commented and coded in my Github under the notebook "ScriptAnalysisFunction.R". The script takes every move from every game and breaks it down from every take / capture that occurred to the "special" moves of chess broken down into more standard move notation, such as castling and *en passant*. For every game, given the script, a list of items were taken out for every piece on both sides of the board: the number of pieces taken, the value of the pieces taken (based on standard chess value), the list of pieces taken by that piece, the number of times the piece moved, and the number of checks the piece made ("attacks" on the king). This was done for every game in the dataset, and is by far the most computationally expensive part of my analysis. Therefore, for simplicity, that is the only difference between "chess_games.csv", the raw data file from Lichess and Kaggle, and "chess_games_analyzed.csv", which has all of these features as separate columns.

Models

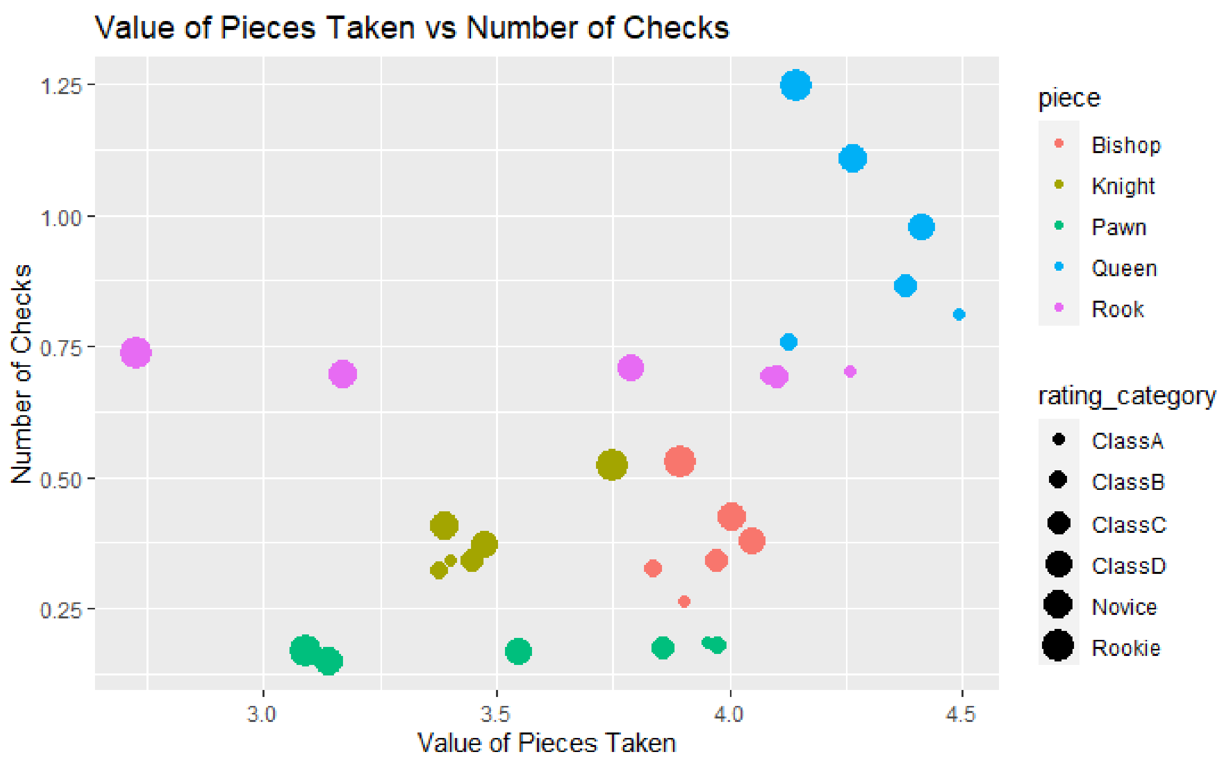
Before any of the refined models, I chose to look into some of the categories by themselves to see if there could be any noticeable differences and trends just from the categories. One of the factors that I did not use in any of the models, but showed some differences among categories is the number of moves per each piece. When looking at the proportion of moves made per piece, we see an interestingly linear difference when looking at just the rook and the queen. As displayed in Appendix Figure 1, worse players used the rook and queen equally, at about 13-14% of total moves for both pieces, while better players, the difference was about 5-6% with the queen being used only about 12% of the moves whereas rooks were about 18%. This does not mean too much to the value of the pieces per say but can tell us maybe a bit about lost value that younger players lose by not using their rooks as often as they should.

Model 1: Takes and Checks Model

For our first model, I looked into creating a model as "naive" as it comes when it comes to evaluating a game and the pieces using two of the features above, the value of the pieces taken by a certain piece for a game, and the number of checks the piece had in a game. I was very much inspired by the idea of "counting stats" in sports, such as Rebounds or Steals in basketball - very interpretable stats, but on their own, do not say much to the outcome of a game. The key reason I chose these two, was they were the simplest "counting stats" that could ideally encompass the entire game in the simplest way possible: the value of pieces taken would be track the entire game's progress and interactions with the opponent's pieces whereas the checks would be the most objective based counting stat that can create a proxy of how well the pieces help achieve the end goal of the game, a "checkmate".

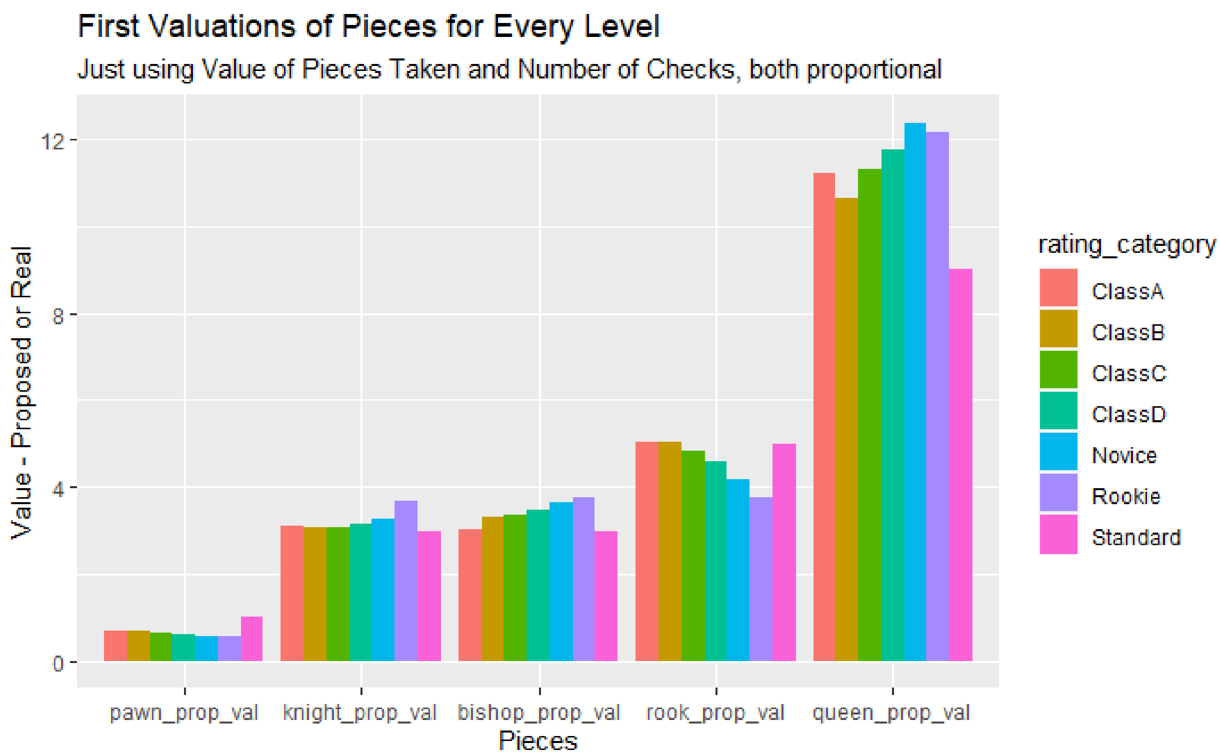
When looking at just the value of the pieces taken by itself, we see some interesting trends across levels. As detailed in Appendix Figures 2.1, 2.2, and 2.3, we see the average value of the pieces taken for each minor piece (knight, bishop, rook) per each level. For each graph, there is a horizontal line across the graphs which indicate the traditional value of the piece (bishop and knight is 3, rook is 5). By this metric alone, we see the Knight overperform its value at every level, at a consistent rate, at about 3.4 for each rating category. The bishop surpassed this value even more, also at a consistent rate at about 3.9. Lastly, the rook sevely underperforms its traditional value, but this is not as consistent across levels. For better players, this value is at about 4-4.25, around a point less than the traditional 5 value, while for worse players, the average value taken dips to about a range between 2.75-3.75, which is equivalent or even less than the bishop and knight. This seems to correspond pretty consistently with the move counter in Figure 1, but also more of a general sentiment shown across all levels: rooks are not meant to take other pieces as much as bishops and knights are because of the timing of their presence in the game. One of the obvious reasons for this is because the rooks start in the corner of the board, and much of the traditional early game strategies stresses fighting for the middle of the board, making the rook a piece not usually involved in the first moves, where potentially some or even most takes happen. Regardless, it does provide an interesting element to our understanding of its value, and if it should be considered so highly compared to the bishop and knight if it is not a key piece in at least a third of most games. But where the rook can thrive is with the end goal of the game, such as by trapping the opponent king in a checkmate, another reason why checks were included in this model

So for the construction of the first model, a key part of it was finding a way to balance the counts of values of pieces taken to number of checks, as the values have a much different variance because of the nature of the actions. As displayed below, we see a graphical perspective of this variance difference among the values for each piece among the rating categories. Within this graph, the larger the dot, the worse the player category, and the different colors correspond to a different piece.



While they are interestingly clustered, there are some key trends, most clearly, the rook (pink/purple dots), has the average number of checks being consistent among players at about 0.75, but this variance of value of pieces taken is still apparent as mentioned before. This is, however, the 2nd highest among pieces, assisting our earlier hypothesis that adding this data point would assist its value for the rook because the stat is more objective oriented. The biggest variance we do see with the number of checks is with the queen and bishop, where better players do not actually get as many checks with these pieces, most likely because there are less opportunities for it against a better opponent.

Given this understanding of these values, and noticing that the number of checks has a range of only 1 total when it comes to averages (~0.25 to 1.25), instead of creating a simple summation like originally planned, I decided to use the proportions of these stats on average. This enabled an equal weighting of the two stats, which do not have any solid justification besides the fact that they encompass parts of the game easily (hence, making this the more naive model of the two in an encompassing naive evaluation of chess pieces). So, once I took the proportions of the value of pieces taken and the number of checks, I added those together and scaled it to 39, the total value of pieces for a certain side in a game. Given these, I found the estimated value for each piece at each level. It is the following:



	rating_category	pawn_prop_val	knight_prop_val	bishop_prop_val	rook_prop_val	queen_prop_val
1	ClassA	0.68	3.11	3.02	5.05	11.23
2	ClassB	0.69	3.07	3.32	5.02	10.63
3	ClassC	0.65	3.08	3.34	4.81	11.29
4	ClassD	0.61	3.15	3.47	4.56	11.77
5	Novice	0.56	3.26	3.66	4.16	12.38
6	Rookie	0.56	3.66	3.77	3.75	12.17
7	Standard	1.00	3.00	3.00	5.00	9.00

On the graph, we have the pink bar on the right to represent the standard values of the pieces, in order to compare the values. Each piece has an interesting unique trend. Each predicted pawn value has a value less than 0.7, with the lower ranking categories at lower than 0.6. These values make sense considering that the values do not take into a key aspect of the individual pawn: the promotion. In the script function, when the pawn is promoted to another piece, the model does not recognize the future moves of the promotion piece. Therefore, a fair estimation of just the promotion value could be estimated as 0.3-0.4 depending on the level. The knight is interestingly spot on to the standard value for each level outside of the rookie, where it is valued at a 3.66 (<.3 difference from 3 for all other levels). The bishop has a similar value for the top levels, but increases for every rating category worse than Class A, the best category, all the way up to 3.77 for the worst player category. The most interesting stat here is that the bishop significantly out values the knight at every level outside of Class A, which adds to this hypothesis that the standard values are made for better play / experts of the game rather than the more casual players. With the rook, we see a similar truth but with an opposite trend, where the value of 5 is pretty close to the top ranking categories' estimated values, but as we go down the levels, we see a decrease. Up till novice, we still see the rook at a significantly higher value than the knight and bishop, but at rookie, the values equalize. Yet again, this makes sense as rookie players typically do not have a grasp of the relative values at all. Overall, it is key to note that the rook values do not exceed 5 at all, putting this relative value ceiling for a rook = 5 value. Then comes the queen, which clearly is a beneficiary from this model because of its ability to move, to take, and to make checks, which is displayed already in the standard value. But this is made even more evident here with all proposed values being ar higher than a 10.5, 1.5 points higher than the 9 value.

Overall, I did find it very surprising how close these values were to the standard values as a whole. Considering that these are proportions, with no heavy influence of the original piece values here, these seem to be relative weights that can tell most of the story of the value of these pieces, or at least how the standard values came to be. One could make the argument that you can get the standard 1,3,3,5,9 values from just the Class A predicted values above, where to make it simply, the pawn is rounded up to a 1, and then after all the other pieces are rounded to the nearest integer, all the remaining value goes to the queen.

Model 2: LogOdds Model

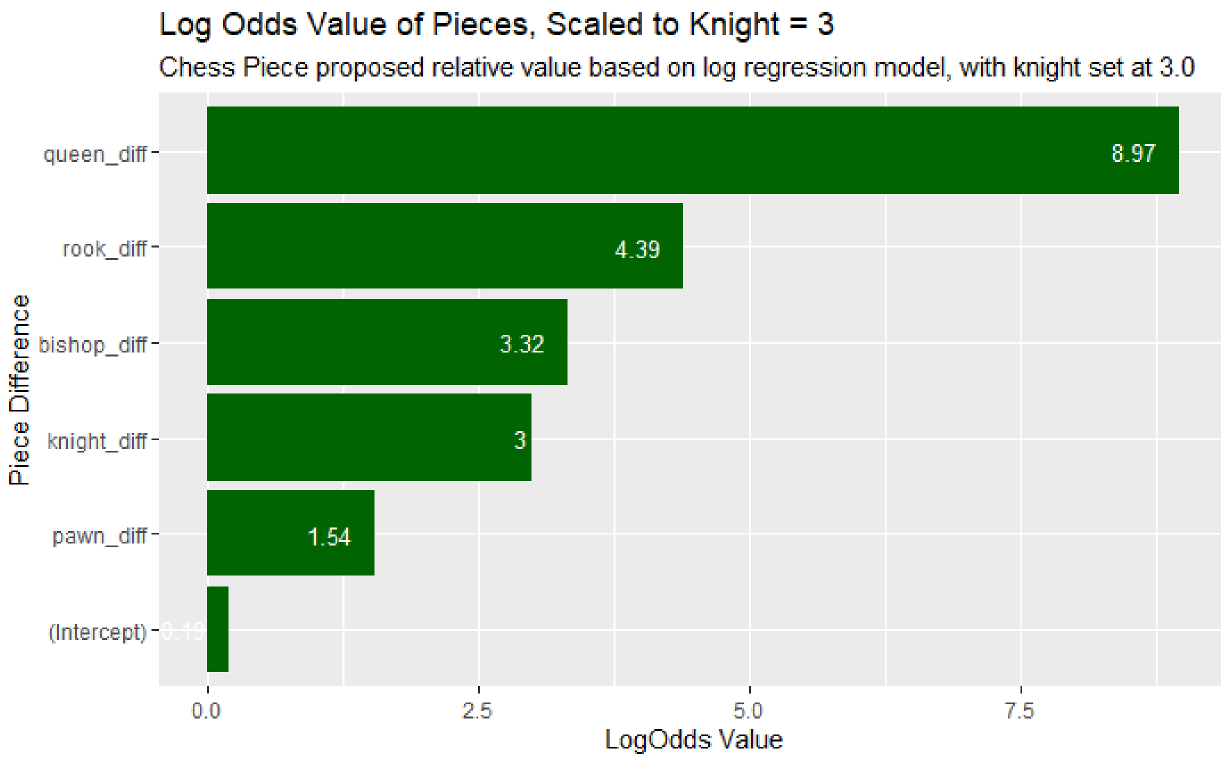
The takes and checks model was great at analyzing all the basic actions that happened during a game and summarizing / averaging those across levels. The one key flaw that it had was its lack of incorporation of which pieces were part of a win. That is what model 2 focuses on.

This model was nearly entirely inspired by Rasmus Bååth, a PHD graduate of Cognitivr Science from Lund Universitywith studies focused on music cognition and rhythm perception. More relevantly for this inspiration, he is an R-blogger on R's blogging site where he posts research studies. In 2015, he analyzed the value of chess pieces via a logarithmic evaluation of the differences of pieces remaining and used log-odds to create a predictive piece value. I did a similar analysis with a couple key differences.

For my model, I got the pieces remaining via the "list of remaining pieces" column created in the script function mentioned above. Subtracting these totals from the original pieces of the game, we get the pieces remaining. This leads to the equation of the logarithmic function:

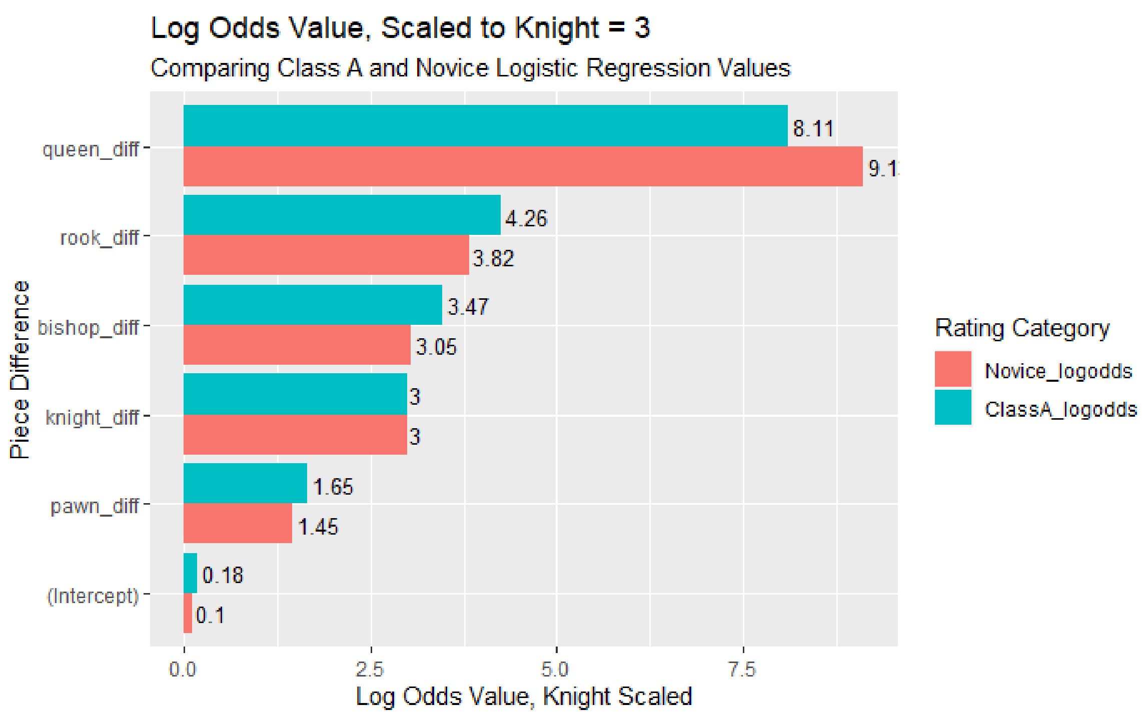
$$winner \sim bishop\_rem + knight\_rem + pawn\_rem + queen\_rem + rook\_rem$$

While this is enough to create relative log odds, these were not too interpretable when we wrote it out. So similar to Baath, I decided to scale the odds by keeping one of the pieces constant to its value. As inspired by Model 1, I chose to keep the knight constant at “Knight = 3.0” because it was most similar across models. This led to the following values:



These values are also interestingly close to the original values, but one key difference from the first model is the pawn, with a value much higher than 1, at 1.54 rather than the low values in the 0.6-0.7 range that were in Model 1. This makes a lot of intuitive sense, because of this promotional evaluation mentioned as a fault of Model 1. Additionally, this articulates a difference with an opponent, and at the end of the game, an additional pawn usually can make all the difference. For the rook, we still see it at a significantly lower value than the expected 5 from the standard values, but still relatively high to the bishop and knight, indicating the value of having an extra rook to your opponent is definitely considerably more valuable than a bishop or knight.

Now when done for all the levels we see the following results. To reduce cluster within the paper, all the individual graphs for Model 2 will be included in the Appendix (Figures 3.1-3.5). If we just compare the Novice and Class A categories side by side we get the following:



Overall, the values show to be pretty similar, with the total net value difference coming from the Novice's Queen at +1, and its losses in the minor pieces: Rook -0.4, Bishop: -0.4, and Pawn -0.2. This overall points to the overuse or heavy reliance of the queen for younger levels, which can be seen mostly across levels as seen below:

	Piece	Novice_logodds	ClassD_logodds	ClassC_logodds	ClassB_logodds	ClassA_logodds
1	Intercept	0.10	0.18	0.21	0.19	0.18
2	PawnValue	1.45	1.51	1.48	1.74	1.65
3	KnightValue	3.00	3.00	3.00	3.00	3.00
4	BishopValue	3.05	3.22	3.48	3.38	3.47
5	RookValue	3.82	4.66	4.61	4.43	4.26
6	QueenValue	9.13	9.30	8.71	7.86	8.11

The slight increase in the pawn value is not statistically significant until we group and compete the bottom 3 to the top 2 levels, where we see a jump in about 0.2. The bishop is always higher than the set knight value besides for the novice interestingly enough. At the top levels, we see nearly a 0.5 difference, which matches a lot of experts' reevaluation recommendation of moving a bishop to 3.5. But this suspected difference is definitely the impact of vision here. Since we are just looking at the end of the game, a bishop has one distinct advantage now with less pieces on the board, and now its the ability to see a maximum of 15 squares in any single position, while the knight is capped at a max of 8. This additional value is key for endgames and most endgame tactics. The rook here shows to have a relatively high value at every level besides the novice, where it lies at 3.82 compared to about 4.5 for the other levels. I was surprised this was not higher considering the commonness of it being used at the end of games, and it points again to this ceiling of the rook's value at nothing higher than a 5. Lastly, the queen, interestingly enough, has smaller than a value of 9 for the top 3 levels. The values are still high and pretty close to 9, but definitely shows the opposite trend compared to Model 1, where the values were much higher than 9.

Overall, this model proved to be fruitful for understanding and very much justifying the value of pieces for different levels of players at the end of the game. For every level besides the beginner, having a bishop proved to be much more valuable than a knight. Alongside that, having a rook and a bishop was just as good as having a queen, for the top 3 levels, but was definitely not as good for the bottom 2 levels. Lastly, having a triplet of pawns can nearly outweigh every piece outside of a queen for an endgame.

Combining Models

Given these models, in order to make an effort to create a summarization of the models, I looked to combine the models to produce a final recommended value for different levels of players. To do this, I went back and forth between deciding if I should weigh them differently or not. I could have weighted the 1st model more than the 2nd because it encompassed more of the game and what occurred, especially when it comes to interactions with your opponent. I also could have weighted the 2nd model more than the 1st because of valuing the endgame and end result / objective more. After the back and forth I decided to simply average out the two values for each with no weights and got the following result:

	Piece	ClassA_comb	ClassB_comb	ClassC_comb	ClassD_comb	Novice_comb
1	Pawn	1.165	1.215	1.065	1.060	1.005
2	Knight	3.055	3.035	3.040	3.075	3.130
3	Bishop	3.245	3.350	3.410	3.345	3.355
4	Rook	4.655	4.725	4.710	4.610	3.990
5	Queen	9.670	9.245	10.000	10.535	10.755

There is a lot to see here, but the one key setback from this average is that half of the knight's value comes from the set knight = 3.0 value we needed for our log odds function. However from my analysis, and in general since we have seen the most consistency with the knight's value across models, it is safe to use that as a given, with the understanding that they are all a bit above 3. My model does therefore imply that for a pawn = 1 valuation, such as the Novice's, 3 pawns would not be enough to cover the value of a night, whereas for Class A, it definitely would. As for other major differences, we see the value of the Pawn for Class A and B as much higher than the lower rated players, maybe because their use and positioning of them is much stronger, especially when looking at all pawns together. The Bishop is relatively consistent among the levels, with Class C having the highest evaluation at about 3.4, and Class A having the lowest value at 3.25. The key theme across is that they are valued about .2-.4 higher than the knight, which is not as significant as the 0.5 difference that other chess studies propose, but do show some significant differences. Future studies should definitely look into replicating the models with paired pieces, and see how much the value of a paired bishop compares to the value of a bishop and a knight, for example, or maybe even a bishop and a rook. With the rook, it is probably the most surprising outcome of this report, with how low the valuations put it at in both scenarios. The utilization and basic traits of the rook are taught and encouraged to master for young chess players, but not as much as the knight and the bishop. It should be noted that arguably the most powerful asset of the rook is its vision, but maybe even more critically the squares it can cut off for the king, or other pieces. Adding a value of piece sight into this model in some way may end up giving the most complete version of the model. Lastly, we see the queen trends we have seen throughout the report with the value falling as the player gets better. For the novice value of the queen to be near 11, it means that it has a higher value than the rook, knight, and bishop values added together.

Conclusion

Overall, the key hypothesis that our current values and understanding of chess pieces may be more geared towards the better players shows to be mostly true. However, there definitely is a lot left to look at here and I want this paper to be a springboard for future analysis and research. GitHub should have everything needed for anyone to replicate the study as needed.

A couple key future parts that I recommend for the future is the following: a) add the categorization of different time intervals, just as was done for rating categories and b) including pairs of pieces, so we can compare and add synergy, and see if there is certain synergies that worse players do not have among their pieces.

A Better Way to Learn Chess

Going back to our initial question about the current ways to learn chess and how it's geared towards perfection and memorization, the biggest next goal of this project is to create a system that opposes this way of learning, and add a new way to better the fundamentals and develop a better understanding of one's personal strengths and weaknesses. Creating a program that takes in one's personal games via the chess scripts, analyzes the games on the piece by piece basis - segmenting them throughout based on a variety of metrics, and then outputting recommendations of ways to improve, such as recommending playing with an opening that gets a rook more involved or an opening that does not rely on the queen to create space and attack as much. This will create a new way to learn and get better at chess that focuses on their own games and how they can fundamentally improve.

References

Chess.com. “7 Tips To Get Better At Chess.” *Chess.com*, 17 March 2022, https://www.chess.com/article/view/how-to-get-better-at-chess. Accessed 13 May 2022.

Chess.com. “Stockfish - Chess Engines.” *Chess.com*, 2020, https://www.chess.com/terms/stockfish-chess-engine. Accessed 13 May 2022.

History.com Editors. “Deep Blue defeats Garry Kasparov in chess match.” *HISTORY*, A&E Television Networks, 16 November 2009, https://www.history.com/this-day-in-history/deep-blue-defeats-garry-kasparov-in-chess-match. Accessed 13 May 2022.

Kaufman, Larry. “The Evaluation of Material Imbalances.” *Chess.com*, 17 November 2008, https://www.chess.com/article/view/the-evaluation-of-material-imbalances-by-im-larry-kaufman. Accessed 13 May 2022.

Rozman, Levy. “How AlphaZero Completely CRUSHED Stockfish.” *YouTube*, 13 May 2021, https://www.youtube.com/watch?v=8dT6CR9_6l4. Accessed 13 May 2022.

Serper, Gregory. “Inexperienced Player Mistakes.” *Chess.com*, 22 November 2015, https://www.chess.com/article/view/the-typical-mistakes-made-by-inexperienced-chess-players. Accessed 13 May 2022.

Staunton, Howard. *The Chess-player's Handbook: A Popular and Scientific Introduction to the Game of Chess, Exemplified in Games Actually Played by the Greatest Masters, and Illustrated by Numerous Diagrams of Original and Remarkable Positions*. Henry G. Bohn, 1847. *Google Books*, https://books.google.fr/books?id=Zu0IAAAAQAAJ&printsec=frontcover#v=onepage&q&f=false.

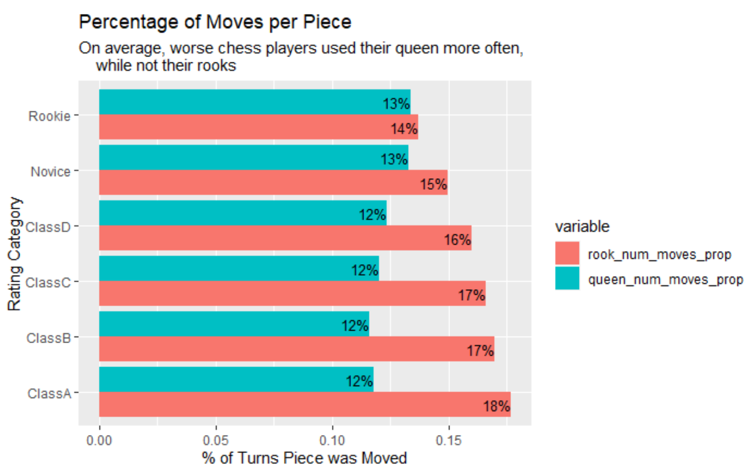
Tomašev, Nenad, et al. *Assessing Game Balance with AlphaZero: Exploring Alternative Rule Sets in Chess*. arXiv, 2020. *arxiv.org*, https://arxiv.org/pdf/2009.04374.pdf.

Vargas, Michael Stephen. “Are Lichess Ratings Accurate? (Original Research!).” *Wegochess*, 2022, https://wegochess.com/are-lichess-ratings-accurate-original-research/. Accessed 13 May 2022.

,

Appendix

Figure 1:



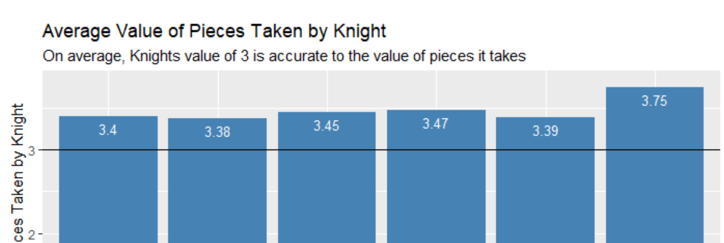


Figure 2.1

Figure 2.2:

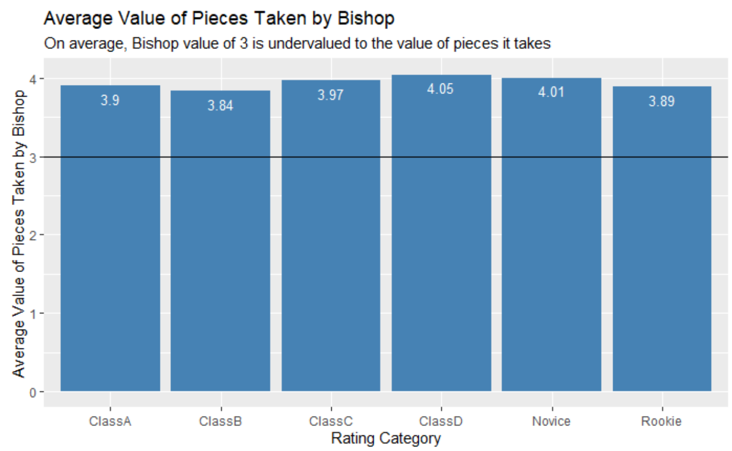


Figure 2.3:

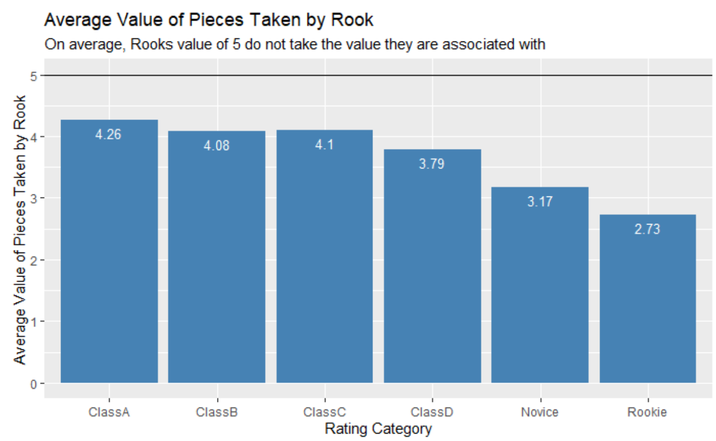


Figure 3.1:

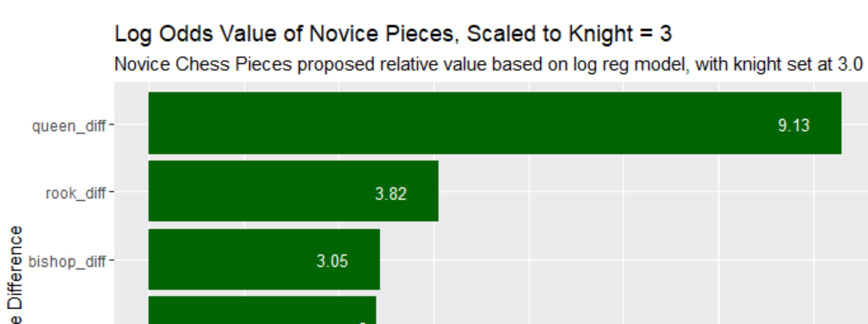


Figure 3.2

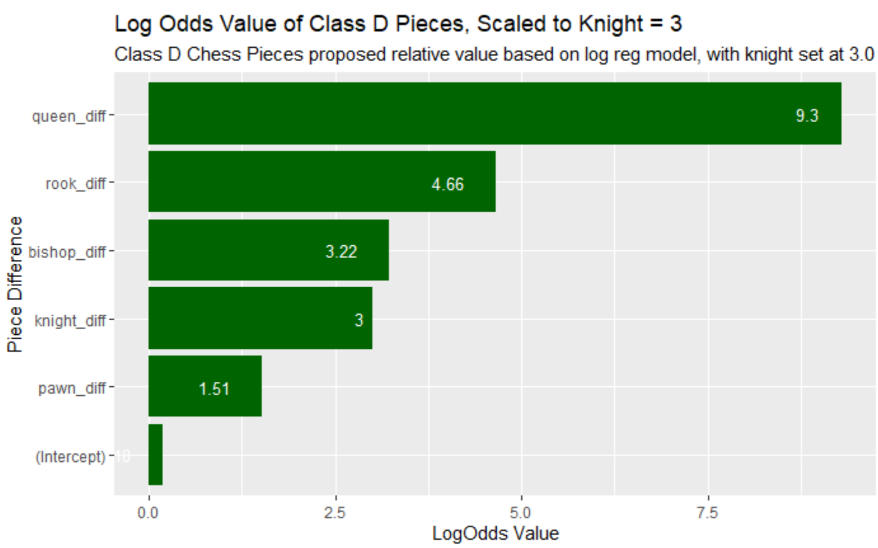


Figure 3.3

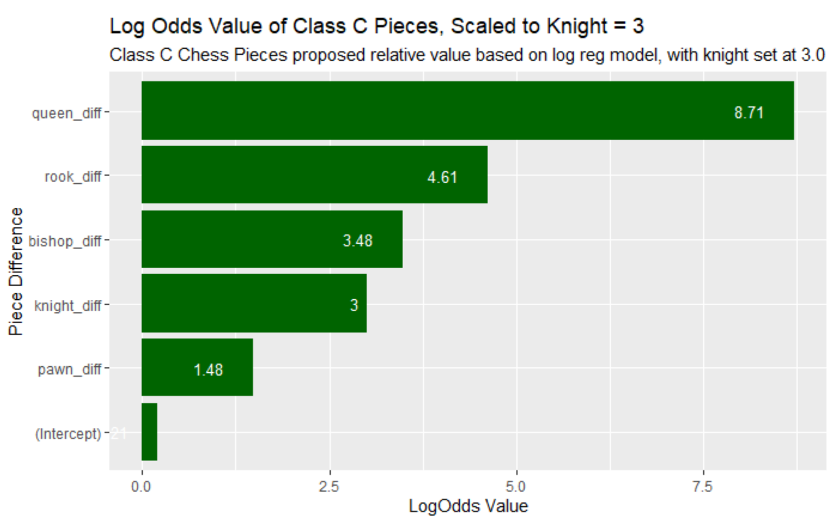


Figure 3.4

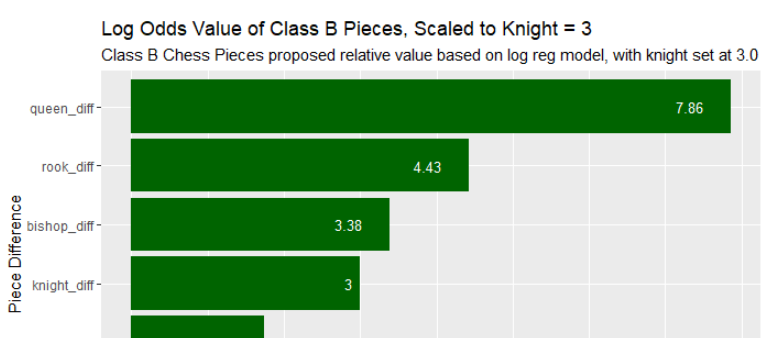
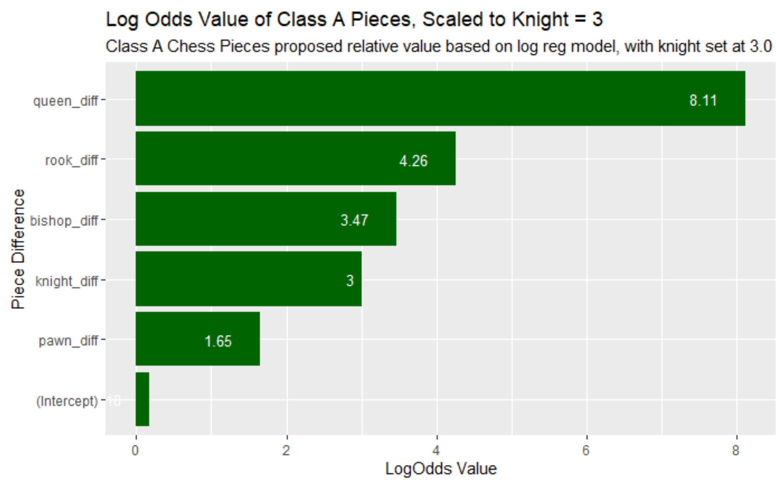


Figure 3.5


Log Odds Value of Class A Pieces, Scaled to Knight = 3
Class A Chess Pieces proposed relative value based on log reg model, with knight set at 3.0
queen_diff
rook_diff
bishop_diff
knight_diff
pawn_diff
(Intercept)
8.11
4.26
3.47
3
1.65
Piece Difference
0
2
4
6
8
LogOdds Value