Drafting *Magic* for Fun, and Maybe Profit

A Dynamic Model, Elo Approach to Draft Analysis

Austin Herrick

April 2, 2022

Abstract

In order to better capture the effects of changing opponent skill throughout a multi-round tournament, I construct a dynamic matching model which pairs players against each other based on record to simulate *Magic: The Gathering* draft tournaments. I approximate the skill level distribution calibrated using win rate data from *17Lands*, and implement an elo rating system to predict match outcomes. I analyze the expected net cost of both Premier and Traditional drafts based on player skill or win rate, along with expected costs of completely collecting a set through drafting. I discuss the ability of each tournament structure to sort players by skill level through matchmaking, as well as the anti-correlation in win or loss streaks that arises from this dynamic matching. I conclude by reconstructing a static model which predicts draft outcomes using only average win rate, and compare these conclusions to dynamic results. I find that record-based matchmaking effectively sorts players against similarly skilled opponents even on short timescales, and that the effects of this sorting on draft outcomes are significant. I find that only approximately 5% of Premier Draft and 7% of Traditional Draft players are able to profit from drafting, and that a more traditional, static model significantly underestimates the net costs of drafts for most players.

Contents

Draft

1 Introduction

The release of *Magic Arena*, and the accompanying ease of drafting at home has led to a dramatic increase in the number of *Magic: The Gathering* players drafting a set many, many times after its release. However, drafts aren't free on *Arena*, and generally cost 1500 gems, or about $7.50[1], each time, an amount that adds up quickly when drafting a set dozens of times. Players can earn those gems back by winning their drafts, but how often does a player need to win to do so? How much do your drafts really cost you, and how skilled would you need to be in order to play for free?

1.1 Previous Answers

There are many ways to answer this question, and several answers emerged when the current draft payouts were released alongside human drafting with *Ikoria* in Spring of 2020. Ryan Spain created a plug-in spreadsheet for players to use[2], and I put together a simple model

[1] At the most generous exchange rate, 20,000 gems can be purchased in the *Magic Arena* client for $100. This works out to 200 gems per $1, or $7.50 for the 1500 gems needed for a draft entry.

[2] See https://tinyurl.com/3b9ts63w

that reached similar conclusions[3]. Those approaches both used the same method – assume a player has an exogenous win rate, and calculate their average prize distribution by flipping coins weighted by that win rate.

But I've never been satisfied by this approach – a player's likelihood of winning isn't static across all their games, because the skill level of their opponents (and the quality of their deck) isn't static either. Pairings in most tournaments, including *Magic Arena* drafts, are based on a player's current record. If you're currently 6-0 in your draft, you'll generally be paired against other players who are 6-0, and those 6-0 players are likely to be more skilled than your earlier opponents (after all, they went 6-0). The reverse is true when you're losing – you'll be paired against other people doing poorly, and your expected win rate will therefore increase, as a losing player is on average less skilled. In general, a player's win rate will converge toward 50% in later rounds of a tournament as they are better matched to opponent's of their same skill level.

2 Using Elo

To better model this situation, we need a simulation with three new properties:

- Players play matches against specific other players, who have their own draft records, and
- Those players need to have a distribution of skill levels, and
- There needs to be a clear way to quantify that skill difference in order to predict a likely winner.

For help, we'll look to chess. Chess uses an elo system to represent the relative skill level of players. Player ratings are centered at 1000, and increase (or decrease) as participants win (or lose) games, with the magnitude of the change depending on the relative difference in elo rankings between players. If Alice wins a game of chess against a player with a much higher elo, her elo will increase a lot more than if she defeated a less skilled player, in recognition of the upset and the higher level of skill it presumably required. Eventually, a player's elo should become relatively stable as it converges to their "true" skill level[4].

In chess, elos tend to follow a logarithmic distribution, or a slightly modified bell curve (with larger extreme tails). Most players are about average in skill level, and higher (and

[3]See `https://mobile.twitter.com/Atnherrick/status/1249794271517040646`

[4]To learn more about the elo system as it applies to chess, Chess.com has a helpful article here[2], and you can read a deeper dive into the mathematics of the system in an article by Jorgen Veisdal here[4]

lower) elos become significantly rarer the farther the are from 1000. For our purposes, I'll be building a similar-looking distribution of *Magic* skill levels using a normal distribution[5].

Because elo is meant to represent the quantifiable difference in skill between two players, it can also quantify the probability of each player's victory in a match, based on their respective elo. In chess, this is done using the below formula:

$$P_A = \frac{1}{1 + 10^{\frac{-(E_A - E_B)}{400}}}$$

Where P_A is the probability of player A winning, and E_A, E_B being player A and B's respective elo scores. As a note, the average elo of 1000, and the 400-point divisor in the formula above, are arbitrary choices. Elo could be centered at any number, and the 400 constant in the above formula only needs to correspond to the update formula - the formula used to adjust a player's elo up or down following a finished match. This choice of numbers has some convenient features though - for example, if Bob has an elo 200 points higher than Carol, he has exactly a 75% chance to win. In our case, we won't be using an update formula, since the elos I assign to players are intended to represent their actual underlying skill level.

Armed with this knowledge, we can estimate the likely outcome of any match of *Magic* based on two player's respective skill levels, and we have an example of what a distribution of player skill levels should look like. All that's left is to determine what *Magic's* skill distribution should look like.

3 Data

To get an idea of how *Magic* skill level is distributed, we'll look to *17Lands*, which has been compiling detailed draft statistics from *Magic Arena* for the last few years. While the website is best known for detailed breakdowns of individual card win rates, their publicly available datasets[6] contain detailed information on players themselves, including how many games they've played, and most importantly, their overall win rate.

For this project, I retrieved data from Best of Three drafts (Traditional Draft) for

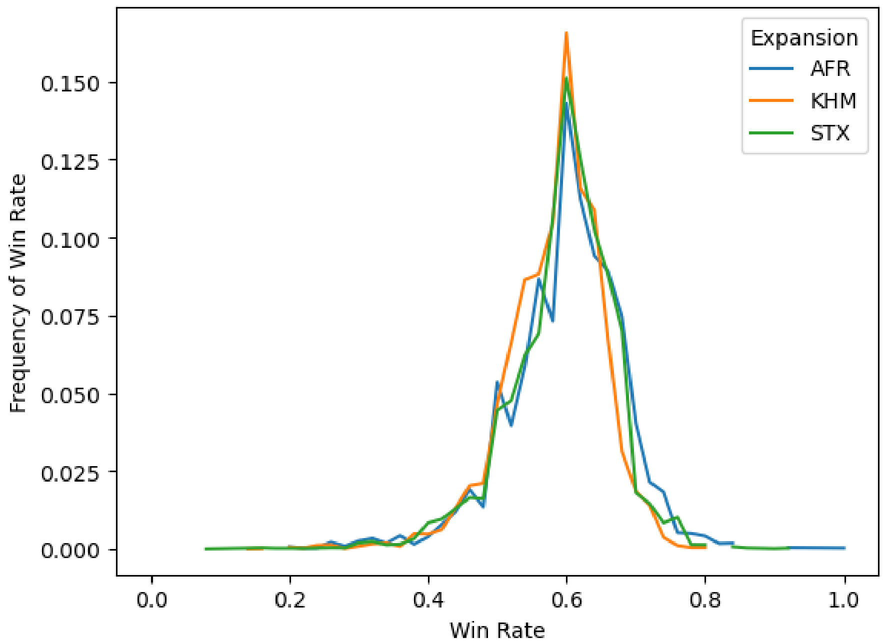
[5] An aside about distributions - Chess's elo system more closely resembles a logarithmic distribution because of its wide tails - a grandmaster level chess player will never lose to a novice, and in fact will very rarely lose to even a highly skilled player, because chess is completely deterministic. In contrast, games of *Magic: The Gathering* involve high amounts of variance - a Hall of Fame *Magic* player can still lose to an inexperienced player due to a poor matchup or a bad draw. Because of this, I use a normal distribution, which has narrower tails than the logarithmic, to better "bound" the possible skill levels. As discussed later, this is also why the likelihood of a player winning a match is capped at 85%, regardless of skill difference.

[6] These datasets can be found here: https://www.17lands.com/public_datasets[1]

Kaldheim, *Strixhaven*, and *Adventures in Forgotten Realms*[7], and Best of One (Premier Draft) drafts for *Kaldheim*, *Strixhaven*, *Adventures in Forgotten Realms*, *Innistrad: Midnight Hunt*, and *Innistrad: Crimson Vow*.

After filtering the data to remove repeated observations of the same draft, we can examine the distribution of game win rates among Best of Three drafters in these three sets in Figure 1 below.

Figure 1: Win Rate Frequency by Expansion, Traditional Draft



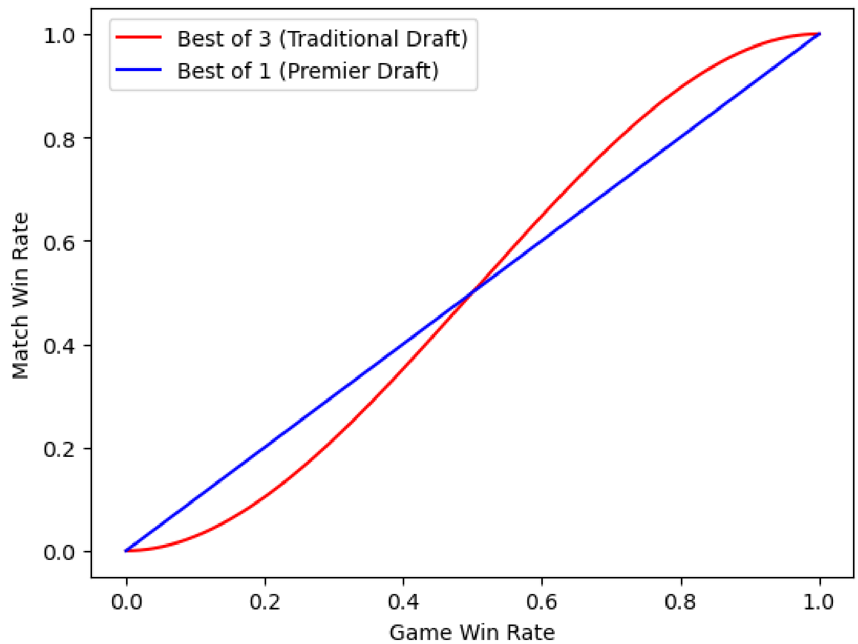
From this figure, we can see that the win rates reproduce the expected pattern. The distribution of elo here is narrower than what we see in chess, which makes sense - chess is deterministic, with no chance elements, whereas *Magic* involves some amount of luck, so a better player might still lose against a much less skilled opponent. There's also a problem though - the average win rate is significantly higher than 50% (specifically, these three sets have a weighted average win rate of 58.9%)! Because *17Lands* only collects data from players using their *Magic Arena* extension, and more enfranchised and skilled players are much more likely to be aware of and interested in such an extension, their data is skewed significantly toward higher skilled players. In my simulation, I assume the true average is 50%, as expected[8].

[7] Traditional Draft data was not available for either Innistrad Set at time of retrieval.

[8] Contrary to intuition, it's not actually required that the average player's win rate be 50%, even though it **is** required that games produce exactly one winner, because players can play different numbers of games. As an example, in a client with 3 players, if Alice and Bob both lose one game each to Carol, the average player's win rate is 33%.

There is a second caveat to remain aware of whenever Traditional Drafts are discussed - the data above reflects a player's **game** win rate, not their match win rate. In a setting where players play a single game (as they do in Premier drafts), this is the same measure, but Traditional Drafts use a best of three structure, where the match win goes to the player who reaches two wins first. As a consequence, relative to game win rates, match win rates for Traditional Drafts are farther away from 50%, because repeated games increase the likelihood of the more skilled player winning. Figure 2 below shows the difference in match win rate from playing a set of three games as opposed to one game, given any game win rate.

Figure 2: Difference in Match Win Rate by Skill Level, Best of One vs Best of Three



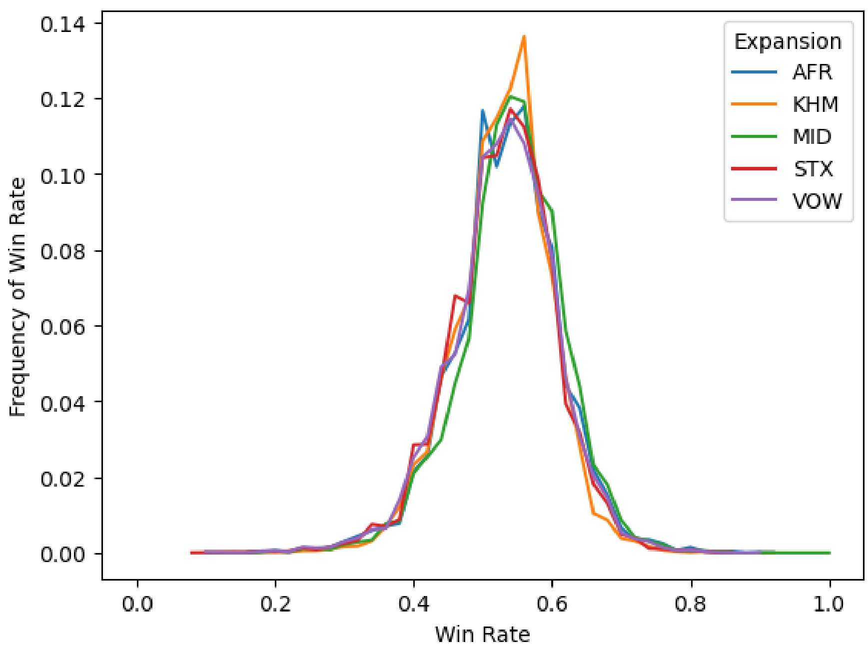
As intuition for this, imagine a match played until a player won 1000 games - while a less skilled player (with, say, a 10% chance of winning) might claim an upset in a single game, they are very unlikely to continue to win against their more skilled opponent over an extremely long series. In particular, a player's likelihood of winning a best of 3 match given a game win rate p_g is shown below:

$$p_m = p_g^2(3 - 2p_g)$$

In general, I will be discussing game win rate in data section (since that is what *17Lands* reports), and match win rate in most other places (since that is the outcome that determines record and prizing). I will endeavour to remove ambiguity between these metrics whenever possible.

Premier Draft data is significantly less skewed in win rate than Traditional Draft data, because Premier Drafts take into account a player's limited rank, which is a loose approximation of elo. This serves the same role as matchmaking in most video games, or elo in chess, by explicitly increasing the likelihood of players competing against similarly-skilled players. For my simulation though, I will be ignoring limited rank when constructing Premier Draft pairings, as simulating the evolution of limited rank, and the pairing priority used by *Wizards* when pairing players, is beyond the scope of this project. Nevertheless, even with this correction, the observed win rate in *17Lands* data for Premier Draft is 53.4%, as seen in Figure 3.

Figure 3: Win Rate Frequency by Expansion, Premier Draft



While I won't be using the average win rate information from these data, there is still a crucial piece of information provided by these distributions - the standard deviation of win rates. Quantifying the variance in skill level between *Magic* players is crucial for parameterizing my model, and we can find the variance of both Premier and Traditional drafts using these sets. I was initially concerned that the standard deviation of win rates might vary significantly between sets. For example, *Adventures in Forgotten Realms* was modelled after a Core Set and had more straightforward gameplay - once a drafter knew that Red-Black was the best color combination, and Blue was very bad, most of the format was figured out. *Kaldheim*, by contrast, was notoriously wordy and complex, and therefore would plausibly give more opportunities for a more skilled player to outwit an average one, increasing the variance in win rates. Conveniently though, this didn't seem to happen. The standard deviations in win rates across sets and format types are reasonably consistent, with an overall standard deviation of 7.49% for Traditional Drafts and 7.6% for Premier Drafts, with format-specific standard deviations shown in Table 1[9].

Table 1: Standard Deviation of Game Win Rate, by Set

Format	**Traditional Draft**	**Premier Draft**
Kaldheim	7.03%	7.01%
Strixhaven	7.62%	7.63%
Adventures in Forgotten Realms	8.34%	7.79%
Innistrad: Midnight Hunt	N/A	7.60%
Innistrad: Crimson Vow	N/A	7.82%

4 Building the Model

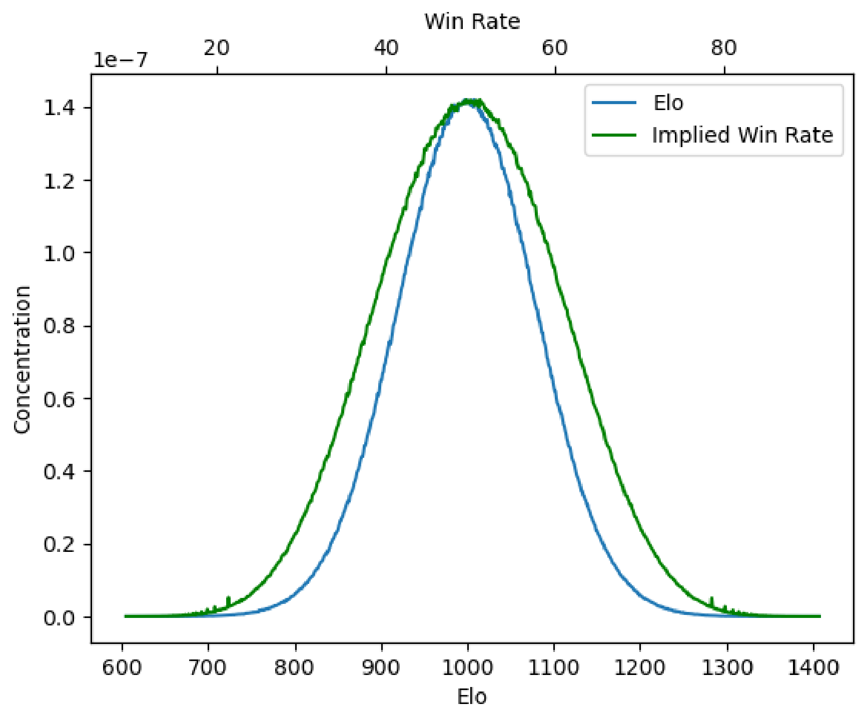
Armed with this information, we're finally ready to construct a set of players with an appropriate and quantifiable distribution of skill levels. To do this, I used Python to generate 100,000 players with elo scores centered at 1000 and distributed according to the standard deviations found in the *17Lands* data. To calibrate the elo scores properly, we need to ensure that a player one standard deviation from the mean has an elo that will predict victory against a 1000 elo player exactly $50+\sigma\%$ of the time, where σ is the standard deviation. We define such a player's elo score as $1000+d$, and solve for d using the elo function discussed above.

[9] The actual model will scale both of these standard deviations up somewhat. Because the *17Lands* data is rightward skewed, it's reasonable to believe that the standard deviation is underestimated. This occurs because the mean of the skewed data is itself close to a standard deviation away from the mean. In a normal distribution, each standard deviation traveled away from the mean reduces the fraction of data contained within one SD relative to the previous SD. For example, the first 3 standard deviations contain 68%, 95%, and 99.7% of the data respectively. This means that data between 1 and 2 standard deviations contains $\frac{95-68}{68}\%$ of the data found within 0 to 1 SD, while data found between 2 - 3 standard deviations contains $\frac{99.7-95}{95}\%$ of the data found within 1 - 2 standard deviations, which is a much smaller fraction. Because this skewed dataset is still calculating a standard deviation to contain approximately 68% of the sample data, we expect a downward skew.

$$0.5 + \sigma = \frac{1}{1 + 10^{\frac{-((1000+d)-1000)}{400}}}$$
$$d = -400 * log(\frac{0.5 - \sigma}{0.5 + \sigma})$$

The above math demonstrates a more general point - an elo score is also a mapping onto, "what is the percent likelihood of winning a game against an average player?". Figure 4 shows the distribution of skill levels generated by my simulation according to both metrics[10] (using the Traditional Draft standard deviation).

Figure 4: Distribution of Elo and Implied Win Rate, Traditional



Before creating the model, there's one more factor impacting win rate we need to discuss - deck quality. Unlike constructed *Magic: The Gathering*, where players can use the same deck for many separate tournaments, drafts requires competitors to construct a new deck every time, which means a player's deck quality can change significantly from draft to draft. Some decks might contain powerful synergies or bombs, while others have mangled curves and inadvisable splashes. For many players, their 7-0 or 3-0 decks don't just represent

[10]For smoothness and clarity, these figures use 10 million players, rather than 100,000

them playing well against worse opponents, but also represent times they piloted unusually high quality decks. This effect pushes in the opposite direction of elo-weighted pairing, by concentrating wins (or losses, for bad decks) sequentially, rather than breaking them up. The "deck quality" parameters discussed later are intended to address this effect.

At this point, we've created a full distribution of skill levels for *Magic* players, and we know the probabilities of match outcome associated with any two of them playing against each other, and we're ready to start our model.

My model consists of a "Tournament Client", which contains all 100,000 players, and simulates thousands of drafts, either Premier or Traditional. A draft functions with the following steps:

1. First, each player is charged 1500 gems and assigned a deck, which has an associated quality, as discussed above. 60% of decks are "average", with quality 0, 20% are "strong", with quality 1, and 20% are "weak", with quality -1.

2. Next, each player is paired at random against another player with the same record. For the first round, all players' records are identical, so pairing is purely random. If an odd number of players exist for a given record, the leftover participant is assigned a hypothetical opponent with 1000 elo.

3. The expected outcome of each match is calculated using each players' respective elo. No player is allowed to have a likelihood of winning a game greater than 85% however, to reflect the fact that any player can draw exceptionally poorly, or mulligan too many times to compete.

4. Next, each match's outcome is weighted by the difference in quality between the two players' decks. A player's chance of winning a match increases by 7.5% for each tier of quality above their opponent's deck (up to 15%)[11].

5. Each match's outcome is determined by chance, using the probabilities arrived at above. In a Traditional Draft, this uses the game win rate to match win rate conversion discussed above. Each player's record of wins and losses is updated, and the next round begins. Steps 2-5 repeat for 3 rounds (in a Traditional Draft), or until all players have at 7 wins or 3 losses (in a Premier Draft).

[11] This is multiplicative, not additive - a player with a 40% chance of winning a match normally whose deck is a tier higher than their opponent will have a final probability of 40 + (40*.075), or 43%, not 47.5%. An additive win probability would mean that better decks were proportionately much more valuable in the hands of less skilled players, while also implying win rates exceeding 100% for players that were already strongly advantaged. It's possible that this assumption could be revisited in further work - I admittedly do not have empirical data to back this up and am instead using my intuition from having drafted hundreds of times.

6. Finally, each player is awarded prizes of packs and gems according to their final record. They are also assigned the rare cards and mythic rare cards we expected them to open in their packs during the draft, for collection tracking purposes.

Note that a player's results in one draft have no impact on their pairings in future drafts, although their lifetime earnings are always tracked. As players accumulate match wins or losses during a draft, their future pairings are more likely to be against players with a similar elo to them, because those players are more likely to have reached similar match results.

5 Validation

The first step in analyzing the results of any model is validating its functionality. To that end, we begin by making sure the emergent behavior of our tournament simulator matches the basic expected results. First, I simulate a single draft using a tournament environment of ten million players, and observe the Elo distribution of players who achieved various records. Figures 6 and 5 show the results for both Traditional and Premier drafts.

Figure 5: Elo Distribution by Win Rate, Traditional Draft

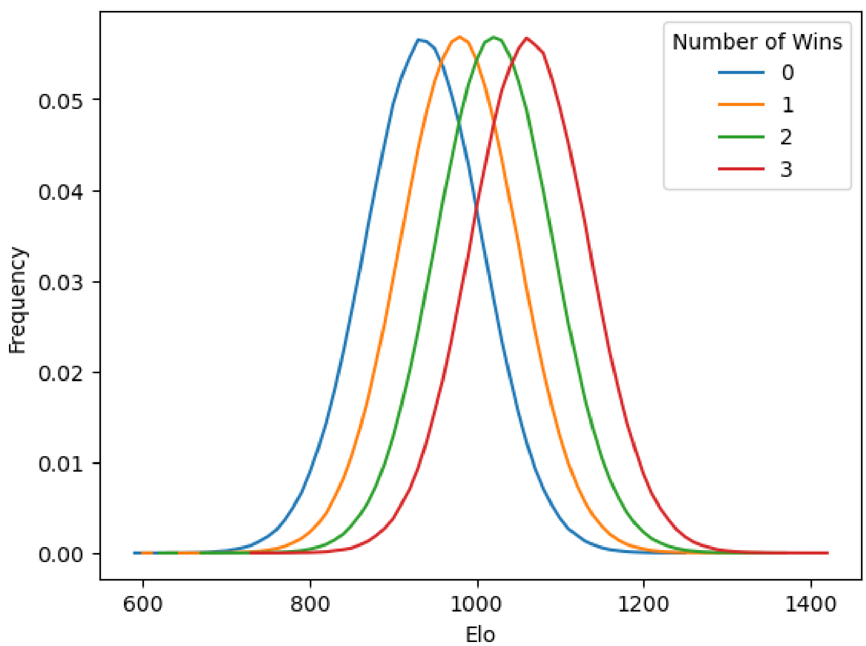
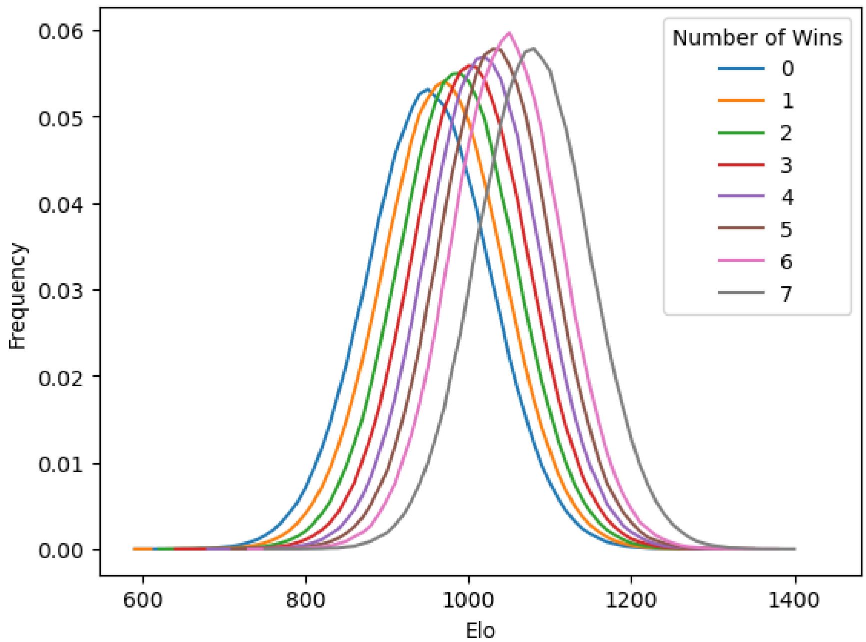


Figure 6: Elo Distribution by Win Rate, Premier Draft



From these figures, we can see how the elo distribution of players shifts throughout a draft as records diverge and more skilled players are more likely to achieve winning records. In Premier drafts, the average elo among 0-3 drafters is 951 (the 27th percentile) compared to 1079 (the 84th percentile) among players achieving 7 wins. For players who reached 7 wins, the average difference is elo between them and their opponent was 83.1 in their first round, versus 19 in their last round.

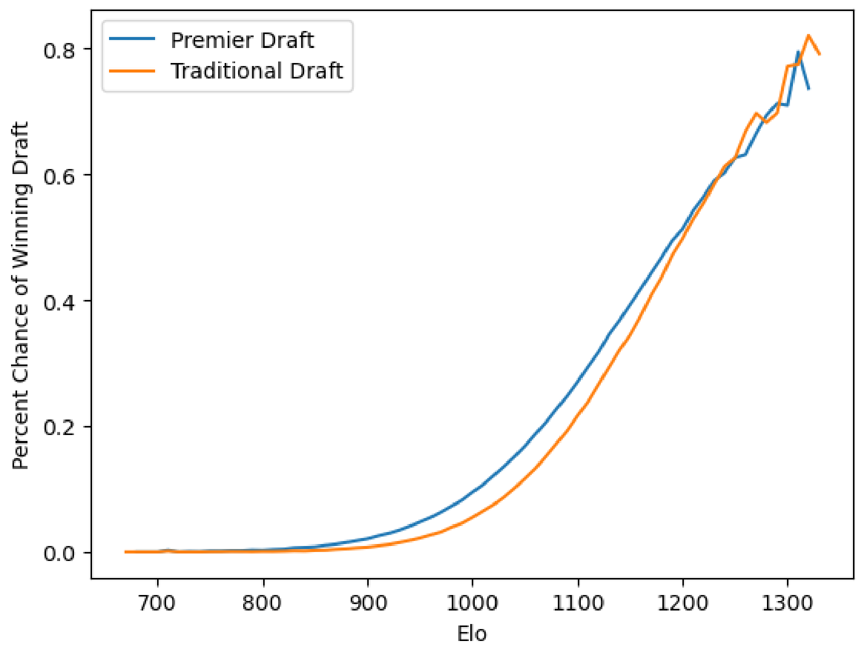
The numbers are slightly different for Traditional Drafts. Traditional Drafts have fewer rounds, which gives less time for elo to fully "sort", but the best of 3 structure means that matches are more likely to identify the more skilled player. The average elo among 0-3 drafters drops to 936 (the 21st percentile), compared to 1063 among 3-0 drafters (the 78th percentile). For players with 3-0 records in Traditional Drafts, the average difference in elo between them and their opponent was 73.2 in their first round, versus 29.8 in their final round. The lower elo among 0-3 players in Traditional Drafts relative to Premier Drafts reflects the impact of the best of 3 structure making worse players more likely to lose, but we can see among drafters who win all of their rounds that the 7-win structure of Premier Drafts has a larger impact than a best of 3 structure in a shorter tournament.

6 Results

6.1 Likelihood of Winning a Draft

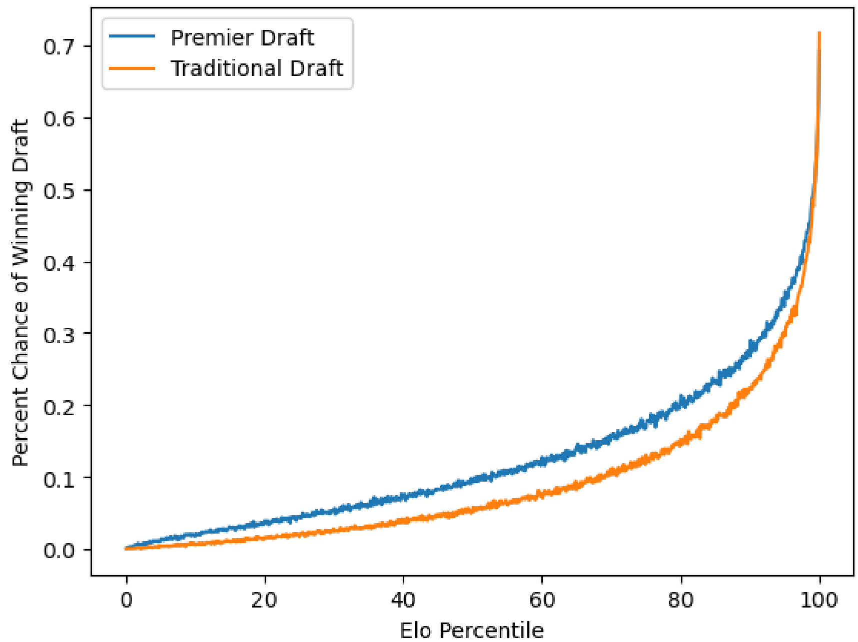
The previous section answered the question, “how skilled are players with various records?”, but we might be more interested in the inverse - how often do players of various skill levels achieve certain records?”. For any level of skill, how likely is a player to ”win” their draft, by achieving the maximum possible number of match wins (commonly referred to as a “trophy”)? To answer that question, Figure 7 below shows the likelihood of players at each elo or win percentage of achieving a 3-0 (in Traditional Draft) or 7 win (in Premier Draft) record.

Figure 7: Likelihood of Winning Draft, by Elo



From the above figure, we can see that achieving maximum wins in a Premier Draft is almost always easier than in a Traditional Draft, until a player reaches an elo of about 1240 (the 99th percentile of skill). Note that the right tail of the graph, where “trophy” rates of greater than 50% are achieved, represents a tiny fraction of players - only 1 in 10,000 players reaches an elo of 1300. Figure 8 recreates this graph with percentile of skill on the X-axis, to make this relationship more intuitive.

Figure 8: Likelihood of Winning Draft, by Skill Percentile



6.2 Implied vs Modeled Win Rates

Earlier, I discussed the concept of implied win rates, or the expected likelihood of players at each elo defeating the average player. But as this article is meant to highlight, drafters do not play against the average player in anything except their first round - instead, they play against a distribution of player skill levels informed by their current record. Because of this, a player's actual win rate will be different than their win rate against the "average" player. In particular, we should expect that the actual win rate will be closer to 50% - record-based pairing operates similarly to an updating chess elo (or video game matchmaking), and will work to pair players against opponents of similar skill levels where the match outcome is closer to a coin flip.

Before visualizing implied and modeled win rates, I should highlight an important distinction - a player's elo score gives us their likelihood of winning a **game** against an average player, but what we're interested in here is their likelihood of winning a **match**, which is a slightly different probability for Traditional Drafts. Because of that, the implied match win rate distributions by elo are not identical for Premier Drafts (which are best of 1) and Traditional Drafts (which are best of 3)[12]. Figure 9 below shows the implied match win rate for players at a given elo percentile for both Traditional and Premier drafts.

[12] These curves are also different due to slight differences in standard deviation, but the differing match structure contributes the vast majority of difference in shape

Figure 9: Implied Match Win Rate, Premier vs Traditional Drafts

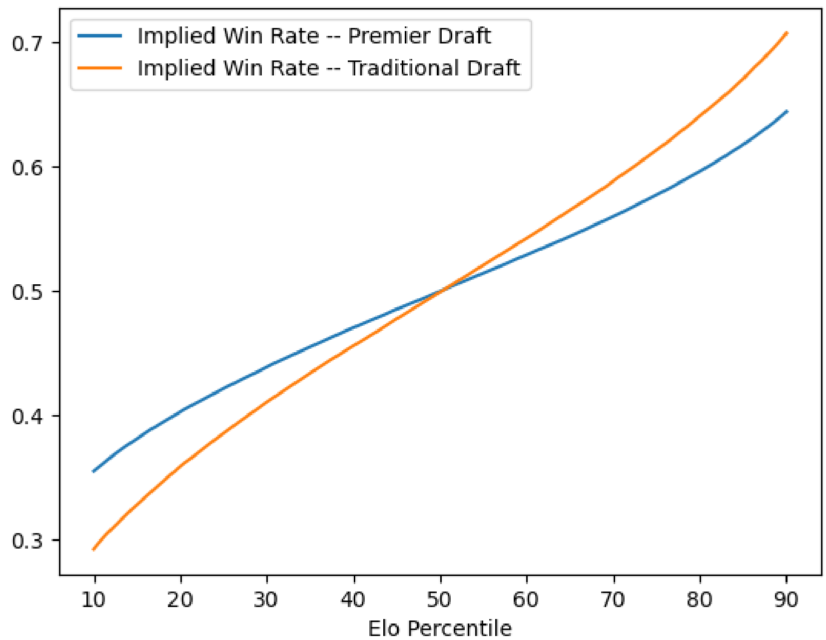


Figure 10 below shows the expected match win rate compared to the actual match win rate for players at each percentile of skill, while Figure 11 shows the difference between the two rates, with positive numbers representing actual performance better than expected performance. As expected, players with below average skill perform better than their skill level would suggest, with the effect reversing for higher skilled players - record based matchmaking makes a difference, even with time frames as short as a draft.

In traditional drafts, this effect is symmetric - the "extra" matches lost by more skilled players are equivalent to the extra matches won by less skilled players. Premier drafts tell a different story though – lower skilled players only win about 1% more matches than expected, but more skilled players lose 3% more. What's happening here?

Figure 10: Implied versus Modeled Win Rate, by Elo Percentile, Premier and Traditional Draft

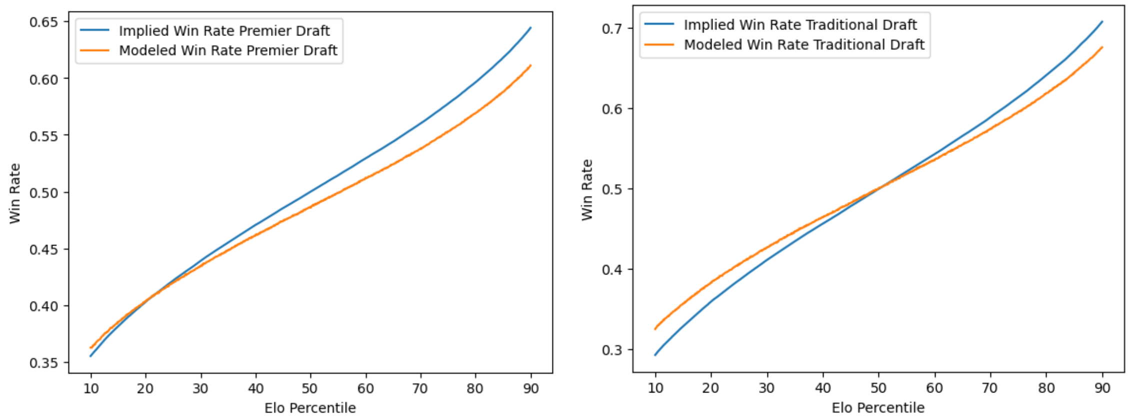
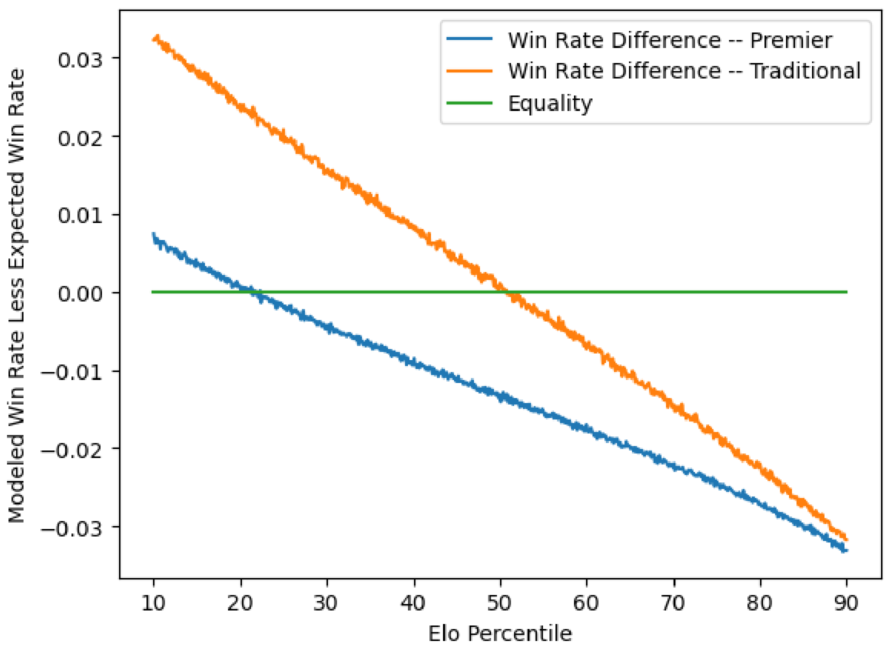


Figure 11: Difference Between Modeled and Expected Win Rate

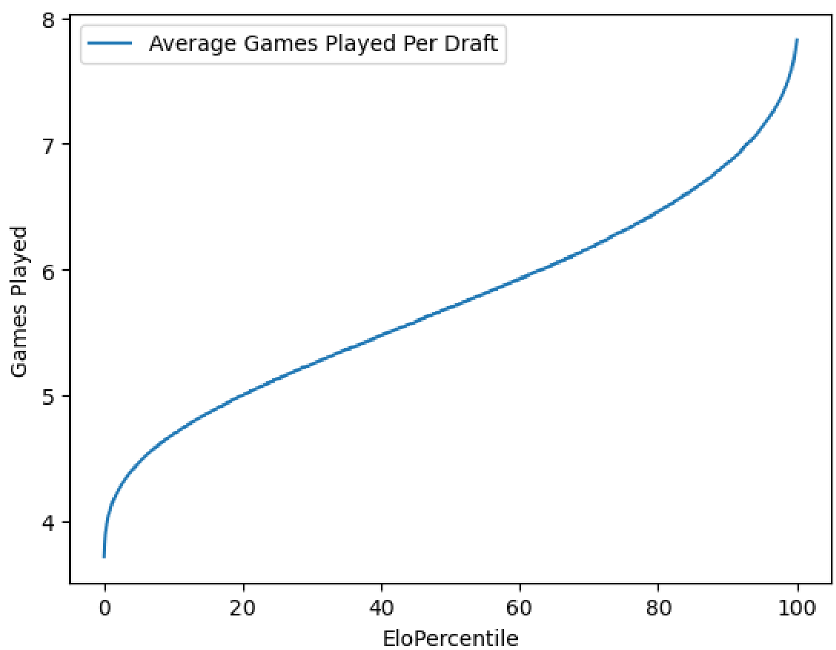


The difference is due to the structure of the tournaments themselves. Players in Traditional Drafts always play exactly 3 rounds, so every player in the client plays an identical number of matches. In Premier Drafts though, players compete until they reach 7 wins or 3 losses, meaning participants play anywhere between 3 and 9 matches per draft. Because winning records involve more games than losing records, more skilled players tend to play more games than less skilled ones. As the draft stretches on, less skilled players reach 3 losses and exit the event, so the only players left are the more skilled ones (or lucky unskilled ones on a winning streak). These players can only possibly play each other, and since every game involves one winner and one loser, this effect drags their win rates back toward 50%.

Another way to think about this effect is to consider that the average win rate in a match is always 50% - one winner (with an 100% win rate) and one loser (with a 0% win rate). If you took 50 players very skilled players made them play some number of matches against each other, the win rate of all players in those matches would inevitably be 50%, and those matches, if counted, would drag down the average participants lifetime win percentage.

I quantify this difference in matches played in Figure 12, which tracks the average number of matches played by elo percentile. As mentioned above, all traditional draft participants play the same number of matches, so they aren't included here.

Figure 12: Average Games per Draft, by Elo Percentile



7 Earning Prizes

7.1 Prize Payouts

Now that we understand how the models works, and what kinds of errors it corrects for, we can finally return to the question that prompted this investigation to begin with - how much does it cost to draft?

The up front cost of both Premier and Traditional Drafts is 1500 gems or 10,000 gold, which works out to about $7.50 (for the purposes of my simulator, all entry fees are paid in gems, and no one earns gold). Once this entry fee is paid, entrants will keep any cards they draft, and will earn packs and/or gems based on their record, with a minimum of 1 pack for all drafters regardless of record. The exact payouts for both draft types, in gems and packs, are listed below.

Table 2: Prize Support, Premier vs Traditional Draft

Table 3: Premier Draft

Game Wins	Gems	Packs
0	50	1
1	100	1
2	250	2
3	1000	2
4	1400	3
5	1600	4
6	1800	5
7	2200	6

Table 4: Traditional Draft

Match Wins	Gems	Packs
0	0	1
1	0	1
2	1000	4
3	3000	6

One complication when evaluating the net cost of a draft is deciding how to value the packs received. Packs are sold within the *Magic Arena* client for 200 gems each, but once purchased, they cannot be converted back into currency or otherwise used to pay for entry fees. There is one exception to this principle - if a user has collected 4 copies of every rare card in the set, any future rare cards opened are converted to 20 gems (Mythic Rare cards are converted to 40 gems once all have been collected). With these factors in mind, we can think of the value of packs falling somewhere in between 20 and 200 gems. In order to allow the reader to make their own decision, Figure 13 below shows the expected net cost of a draft under either assumption, and Figure 14 shows the expected pack rewards. These x-axis on each of these graphs reflect the skill percentile of players, because the expected game win rate (and to a greater extent match win rate) should not be compared on a one to one basis, as Figure 10 showed.

Note that the figures reflect expected game win rate, not match win rate - this allows easier one-to-one comparisons between Premier and Traditional drafts, because any given elo will produce a more extreme match win rate in a Traditional Draft, due to the best of three structure.

Figure 13: Net Gems per Draft, by Elo Percentile

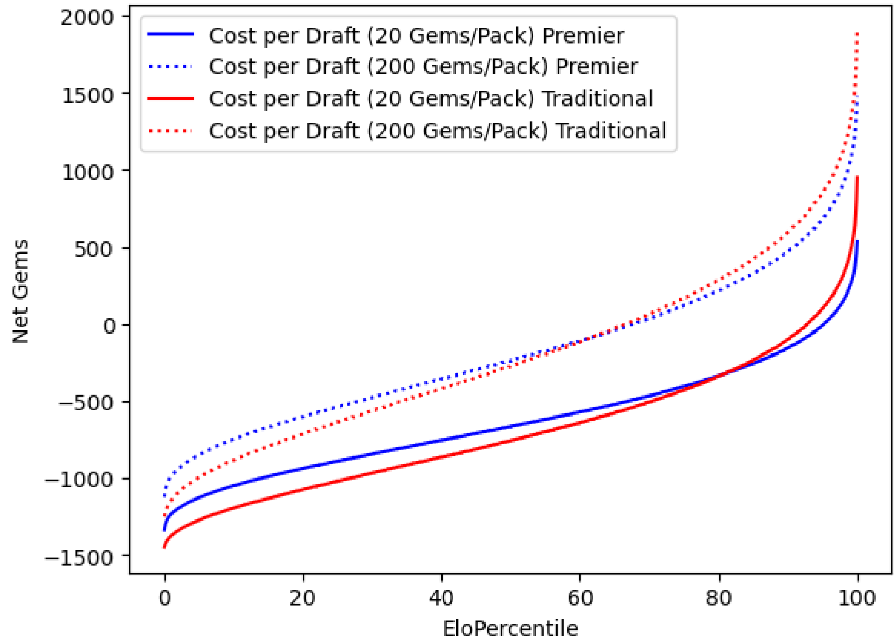
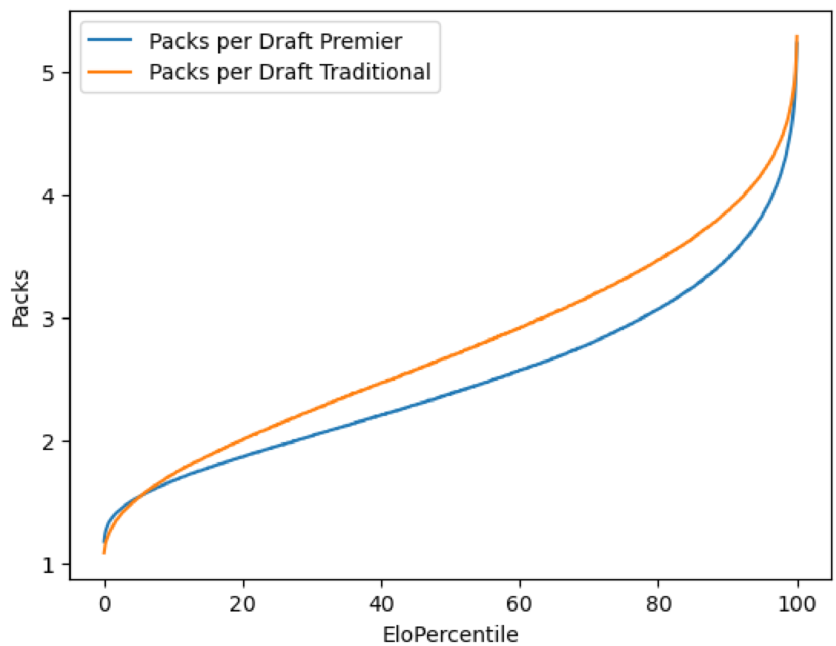


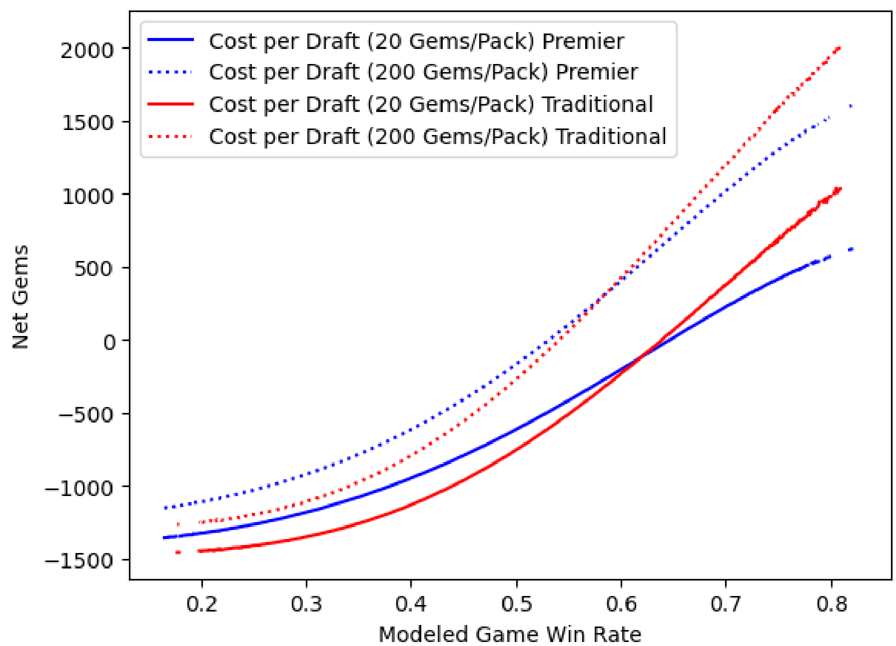
Figure 14: Average Packs per Draft, by Elo Percentile



From the above figures, we can make a few observations. First, at lower skill levels, Premier drafts seem to be a better deal than Traditional Drafts. In particular, for players who win less than half their games, Premier Drafts are generally about 100 gems cheaper if packs are valued at 20 gems, and about 50 gems cheaper if packs are valued at 200 gems instead. At higher win rates though, Traditional Drafts reward more, although when this crossover point occurs depends on how much players value their packs. If packs are valued at 200 gems each, this crossover point happens at the 62nd percentile of elo[13], and if packs are valued at 20 gems each, this occurs at the 80th percentile of elo[14]. The reason this crossover happens earlier with higher pack values is because, as seen in Figure 14, Traditional Drafts are slightly more generous in rewarding packs for almost all players. It's worth reiterating again though that these are theoretical comparisons - my model does **not** take into account limited rank, and therefore assumes a player's average first round opponent's are equally skilled, regardless of their own skill level. A model incorporating rank would likely find Premier Draft rewards for most players clustered more tightly around the 50% win rate rewards, as a player's rank more directly matched them against similarly-skilled drafters.

We can also see how few players are really able to turn a profit drafting - if we assume that packs are worth 20 gems, then players breaking even in Premier Drafts represent the 95.2nd percentile, and those breaking even in Traditional Drafts represent the 92.9th percentile. If packs are worth 200 gems each, the odds look markedly better – players at the 67th percentile break even in Traditional Drafts, compared to players at the 68th percentile for Premier drafts.

Figure 15: Net Gems per Draft, by Modeled Win Rate



For convenience, Figure 15 above shows the expected net cost per draft by game win

[13]This represents a game win rate of 53% in Traditional Drafts, and 52% in Premier Drafts.
[14]This represents a game win rate of 58% in Traditional Drafts, and 57% in Premier Drafts.

rate. As discussed above, the Traditional and Premier draft lines should not be compared directly, as a player at a particular elo will exist on different points along each line.

7.2 Collecting a Set

Another goal players sometimes have while drafting is to collect cards they need for Constructed formats. Because players keep any cards they draft or open in prize packs, drafting can be extremely efficient way to build a collection. One metric of collecting players are frequently interested in is collecting a full set of 4 Rare cards in a set. This is also the point at which packs begin giving players 20 gems rather than rare cards (because they've collected them all), and so is a reasonable stopping point for measuring collection building.

Wizards releases information about the contents of both prize packs and the booster packs used for a normal draft[15]. From this information, we know that the booster packs used for drafts in *Magic Arena* contain a rare card 7/8ths of the time, and a mythic 1/8th of the time (just like Draft Boosters sold in stores). We also know that booster packs awarded as prizes have the same rare/mythic ratio on average (there is some variance from set to set), but 1/30 packs have their rare or mythic slot upgraded to a Wildcard of the corresponding rarity instead. Further, we know that each set has approximately 60 rares found in draft boosters (with some variance from set to set), which means that "collecting the full set of rares" requires 240 rare cards. If we assume that each player in a draft will collect 3 Rare and/or mythic cards per draft (which on average, must be true), we can estimate how many drafts it is likely to take for a player at any given skill level to collect all the rares in a set, and what their net cost for doing so will be.

A brief aside about duplicates - the model described above calculates how long it takes for a player to collect 240 rares, which isn't quite the same thing as "4 copies of 60 different rares". While *Magic Arena* has had duplicate protection for its prize boosters for several years (a user cannot open a rare or mythic rare card they already own 4 copies of), no such protection exists for drafts. My model assumes duplicate protection for draft as well, and while that's not quite true, it's a fairly close approximation, provided a player doesn't open any of their prize packs until doing so will complete their set collection. Combined with the fact that players have agency over which rare cards that pick in drafts, it's likely that duplicates here play only a very small role in changing the time necessary to complete a set. I also won't be counting any rares received through ICRs (Instant Card Rewards) received through daily wins or Mastery Track advancement.

[15] See https://magic.wizards.com/en/mtgarena/drop-rates[3] for a summary of digital drop rates.

Figure 16: Average Drafts to Collect All Rares, by Elo Percentile

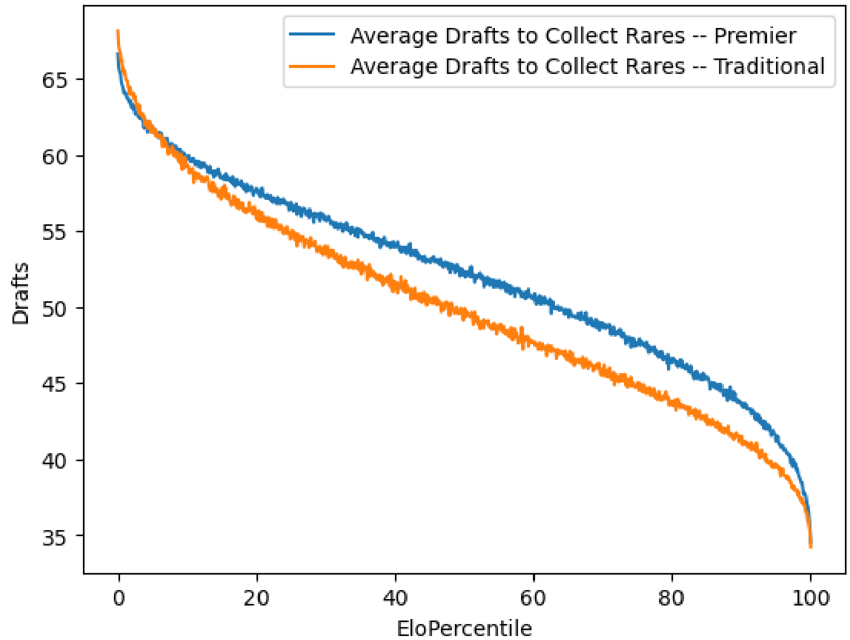
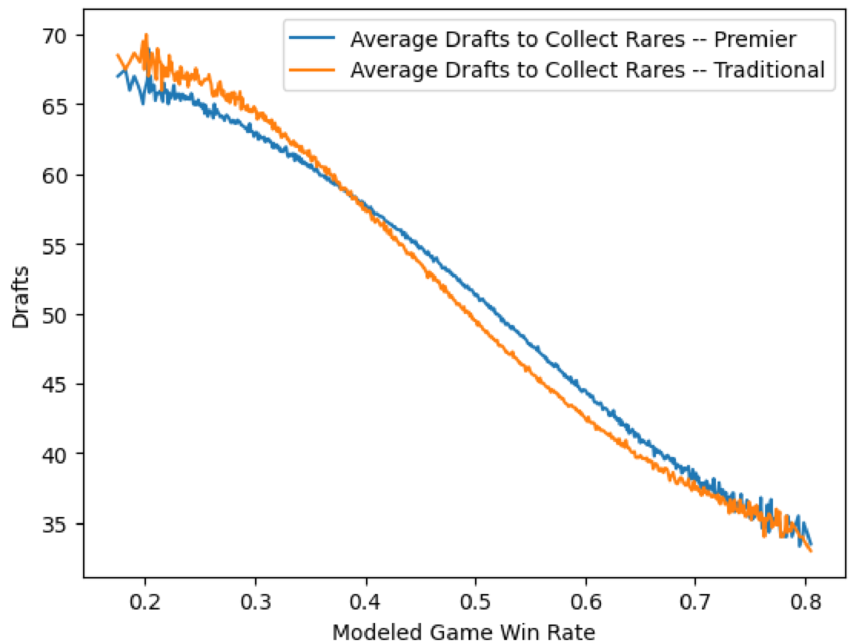


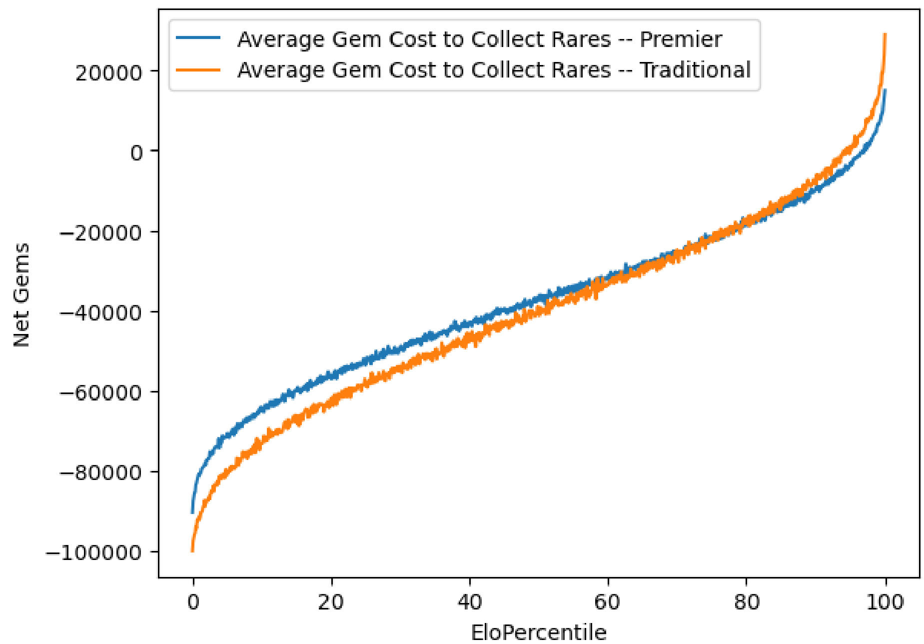
Figure 17: Average Drafts to Collect All Rares, by Modeled Game Win Rate



Figures 16 and 17 above show the expected number of drafts for a player to collect 4 copies of every rare in a set, exclusively by drafting. We can see that win rate affects the number of drafts needed to complete a set in the expected fashion - very skilled players can finish a set in about half the drafts that a novice would need. When looking at elo percentile though, we can see that "around 50 drafts" is a reasonable rule of thumb for how many times the average player should expect to draft a set before collecting all of its rares.

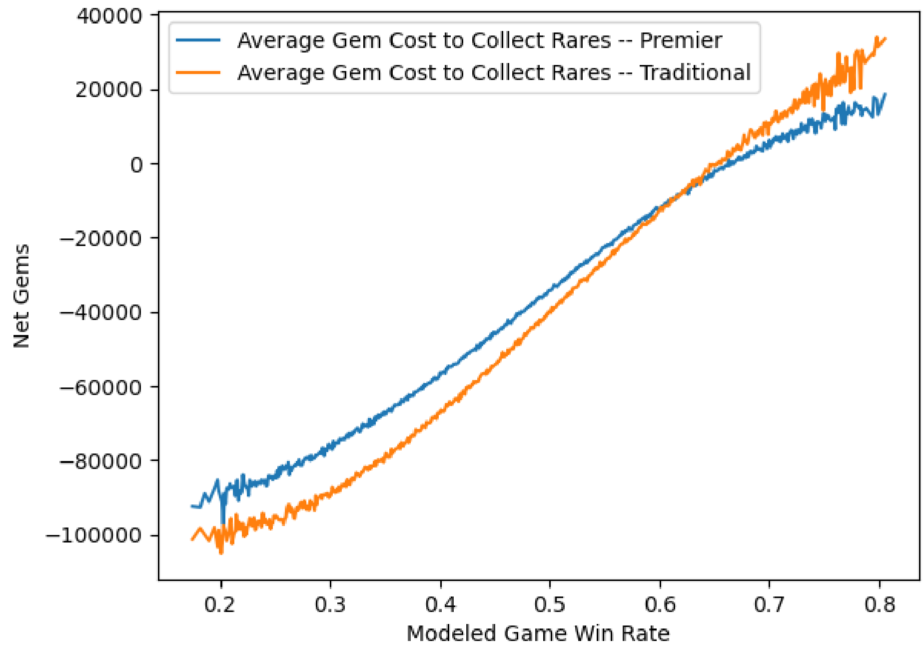
The difference in collection speed is starker when we examine the gem cost, however. Not only are less skilled players entering more drafts to finish their sets, they're paying more for each of those drafts as well. Figures 18 and 19 below show the expected gem cost of finishing a set. For the average player, set completion will ultimately cost 37,500 gems on net, or about $187[16] playing Premier Drafts, and 40,800 gems on net, or about $204, when playing Traditional Drafts.

Figure 18: Average Gem Cost to Collect All Rares, by Elo Percentile



[16] Using the exchange rate of 200 gems for $1, as discussed in the introduction

Figure 19: Average Gem Cost to Collect All Rares, by Modeled Game Win Rate



8 Comparison to a Static Model

The last question we can ask is "how much does this dynamic model matter"? We know that the percent of games a player wins is different from how many games they **would win** against average opponents, but players can't observe their hypothetical skill against an average opponent anyway. Unlike *Magic Online's* formalized rating system, *Magic Arena* does not have a formalized elo, outside its Limited Ranking system, which is different along a number of dimensions[17]. Instead, all players can observe is their realized win rate against actual opponents in their matches, and this statistic is already a product of record-based matchmaking. How much more of the matchmaking effect is left unaccounted for with this simple overall win rate number?

8.1 Trophy Rate

To test this, let's first look at the difference in how often a player will "trophy", or reach the maximum number of wins, in my simulation as compared to applying their modeled match

[17]There's several problems with using Limited Ranking as an elo proxy. It decays significantly every season, forcing players to play a large number of drafts to re-reach a ranking corresponding to their true skill level. At low ranks, it rewards more points for wins than it deducts for losses, meaning that ascension is merely a patience game. Most importantly, below mythic, wins and losses are always symmetric in impact, regardless of an opponent's skill.

win rate a more straightforward "static" model, like the ones discussed in the introduction. In these static models, no dynamic effects exist - players don't compete against other players, they merely apply their win rate to a die roll until the tournament ends, and each round's outcome is independent of the previous one.

In a Traditional Draft, this is simple - a trophy record requires winning three rounds in a row, so the likelihood of a 3-0 record at a given match win rate w_p is w_p^3. In Premier drafts, this is a bit more tricky, since there are several ways to reach 7 wins, because a 7-0, 7-1, and 7-2 record all imply the maximum possible rewards. Luckily, we can re-frame this question by asking, "what is the probability that a player at a win rate w_p will win at least 7 of their next 9 games?"[18]. Combinatorics provides us with a handy formula for this, interpreting each match as a Bernoulli trial -

$$T_r = \sum_{i=7}^{9} \binom{9}{i} (w_p)^i (1 - w_p)^{9-i}$$

Using these formulae, we can replicate the expected trophy rate using only win rate, and compare it to the likelihood of winning a draft from my dynamic model. Figures 20 and 21 below show both statistics for both Premier and Traditional draft, and Figure 22 shows the ratio in actual versus implied trophy rate for both draft types.

[18] This is equivalent, because we can treat 8-1 or 9-0 records as representing 7 wins. Players with those records must have achieved a 7-0 or 7-1 record at some point during the 9 games, "trophying" the draft.

Figure 20: Premier Draft Trophy Rate, Dynamic Model versus Probability Analysis

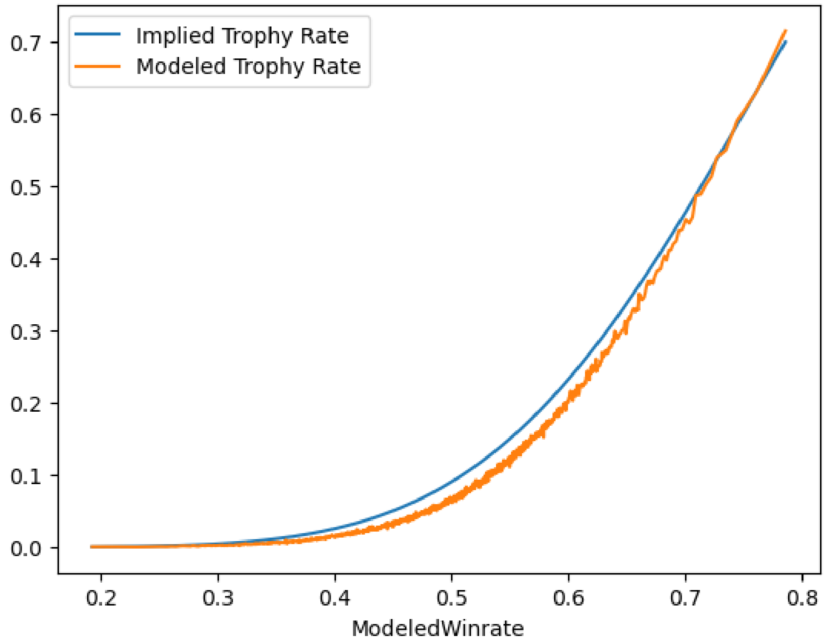
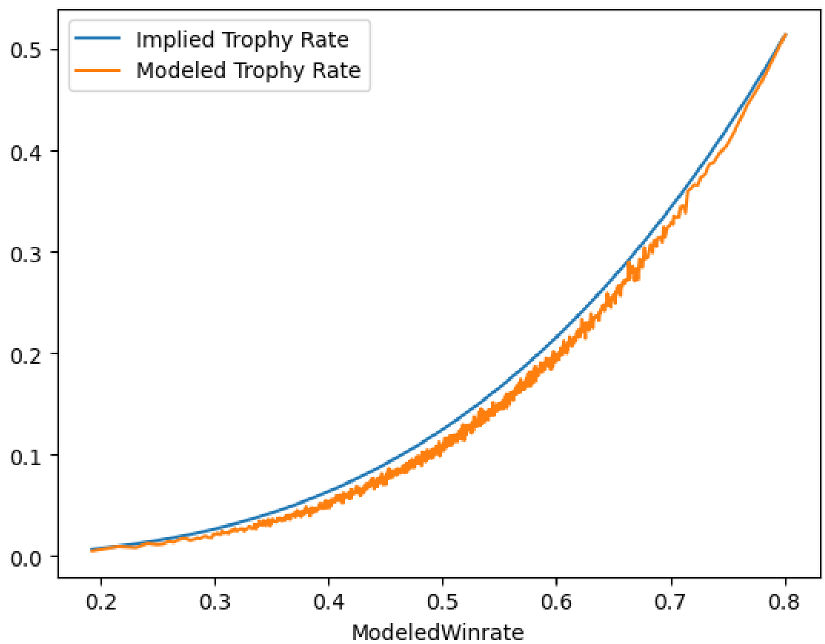
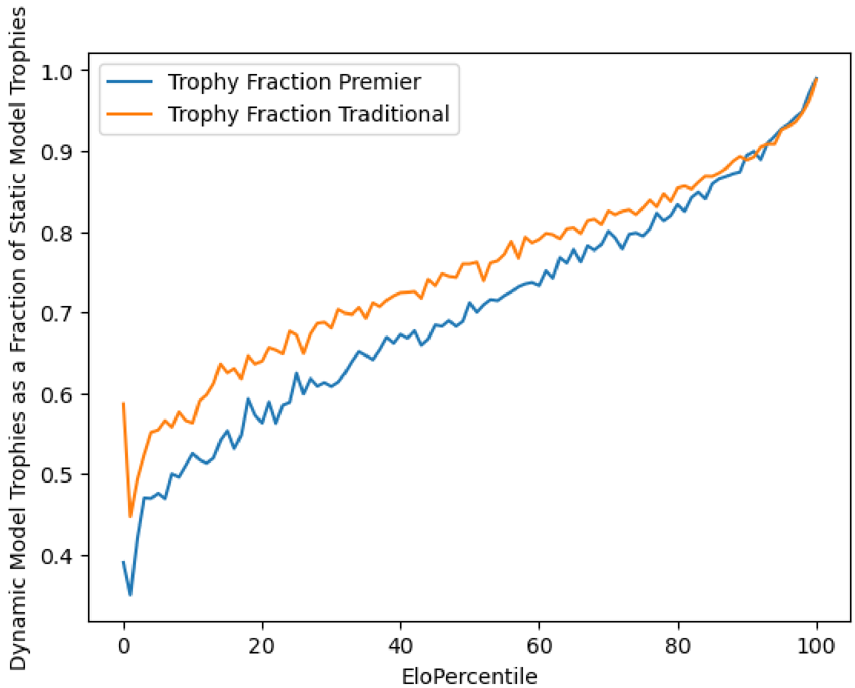


Figure 21: Traditional Draft Trophy Rate, Dynamic Model versus Probability Analysis



Draft

Figure 22: Dynamic Model Trophy Rate as a Fraction of Simple Probability Model Trophy Rate



From these figures, we can see that a significant gap in trophy rate exists, especially for lower skilled players. While the win rate graphs obscure this somewhat (because trophies are relatively rare for most drafters), Figure 22 shows that the winning draft records for the average player only occur 76% as often in Traditional Drafts in my dynamic model as compared to the static model, and only 71% as often in Premier Drafts. For players at the 25th percentile, these gaps are even larger - trophies are 62% (Premier Drafts) or 67% (Traditional Drafts) as common for them.

More generally, the gap between expected and actual trophy rate shrinks as win rate improves - this actually demonstrates the theory discussed in the introduction. The reason the penalty exists at all is because the competitors standing in the way of a perfect record also have good records and are therefore more skilled than a player's average opponent. This causes a player's win rate to drop when they are otherwise very close to reaching a draft trophy. However, very skilled players find themselves in contention to win draft trophies very frequently, so these positive-record opponents **are** common opponents. For that reason, their average win rate is already a reflection of their experience playing against trophy-eligible opponents. The skill advantage of those opponents has already been baked in for very skilled players, and is in fact the source of the difference between implied win rate (based on elo), and empirical actual win rate.

8.2 Expected Value, Static vs Dynamic Models

We can also examine the overall expected net cost per draft suggested by the static model, as compared to the dynamic result we looked at above. To do this, we'll need to flesh out the formulas above in order to capture the likelihood of every possible draft outcome. For Traditional Drafts, this requires calculating the probability of a player at any match win rate, w_p, winning either 0, 1, 2, or 3 matches, multiplied by the prizes for that record. We can sum that total using the formula below, where G_i represents the gem prize associated with an i win record:

$$G_t = \sum_{i=0}^{3} \binom{3}{i} (w_p)^i (1 - w_p)^{3-i} G_i$$

In a Premier Draft, competitors stop playing once they reach 3 losses. Because of this, a "4 win record", for example, requires a player to win exactly 4 of their first 6 games, followed by losing their 7th game (because once they reach their third loss, they are ineligible to play more games). More generally, for any record with i wins, a Premier Draft player must achieve those i wins in $i + 2$ games, followed by a loss - which occurs with probability $(1 - w_p)$. The final formula for the gems earned from Premier Drafts is the below[19]:

$$G_t = \sum_{i=0}^{6} \binom{i+2}{i} (w_p)^i (1 - w_p)^3 G_i + \sum_{i=7}^{9} \binom{9}{i} (w_p)^i (1 - w_p)^{9-i} G_7$$

Using these formulas, we can find the expected net draft cost for each type of draft in the static versus dynamic model based on a player's match win rate. Figures 23 shows the expected net gems per draft (under either pack worth assumption) for Traditional Drafts, and Figure 24 shows the same information for Premier Drafts.

[19]7 win drafts are calculated separately, using the formula from the previous section, because there are multiple possible records that achieve 7 wins, depending on the number of prior losses

Figure 23: Net Gems per Traditional Draft, Dynamic versus Static Model

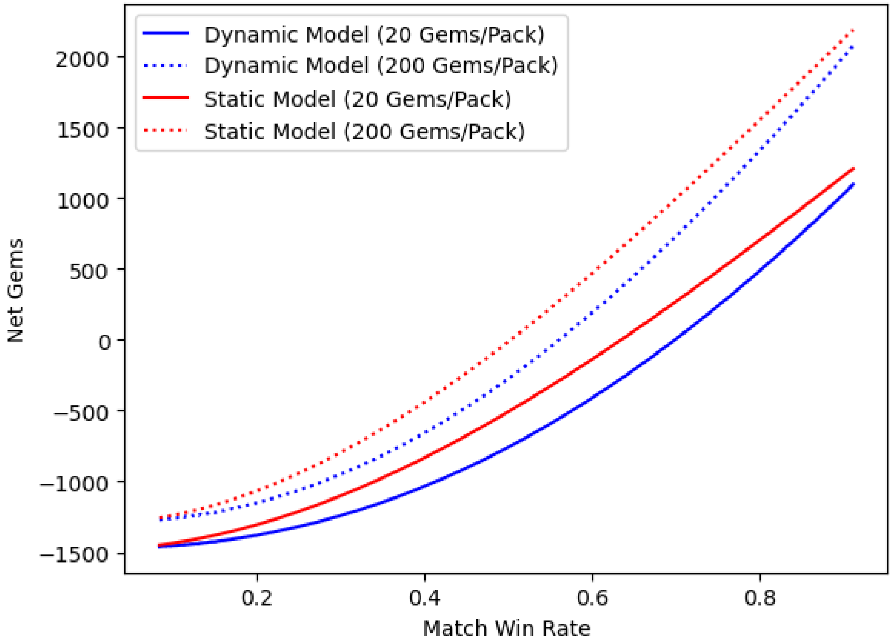
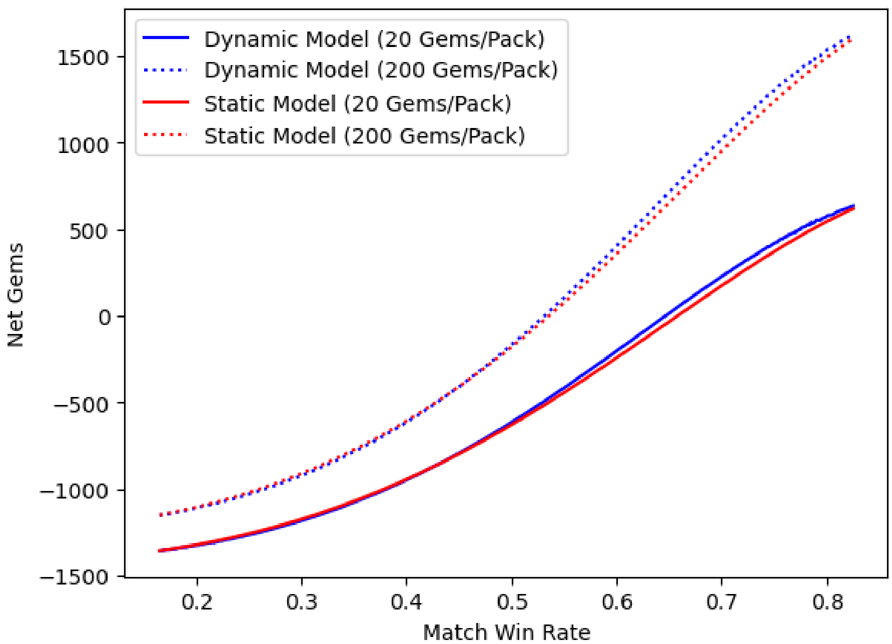
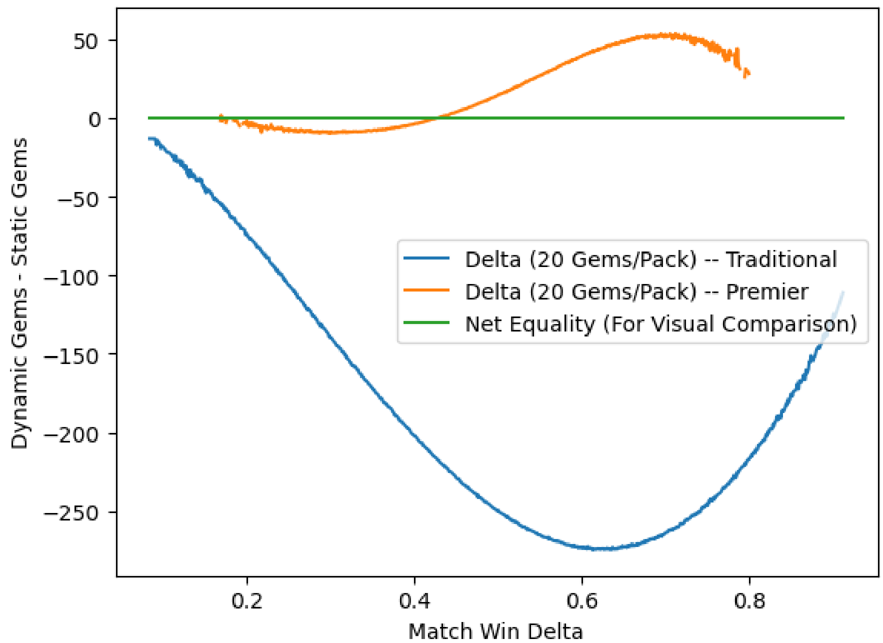


Figure 24: Net Gems per Premier Draft, Dynamic versus Static Model



From the above figures, we can see that the static model suggests a significantly lower net draft cost across all win rates in Traditional Drafts, with the largest difference - over 250 gems - occurring at match win rates of about 60%. In contrast, the static model captures the dynamic results of Premier Drafts very accurately, with only an extremely small marginal difference. Figure 25 clarifies these results by mapping the difference between the dynamic and static models for both draft types by win rate (here, negative numbers indicate that the dynamic model rewarded **fewer** gems than the static model).

Figure 25: Difference in Net Cost per Draft, Dynamic versus Static Model



So, why does the static model do a much better job of approximating the dynamic results for Premier Drafts than Traditional Drafts? There's two reasons - one simple, and one much more complicated. The simple answer is that Premier Draft payouts are much flatter[20] and less "winner-take-all" than Traditional Drafts, so small deviations in draft outcome in either direction make a smaller difference than in Traditional Drafts, where prizes are extremely top-heavy, and players can only turn a profit by achieving a 3-0 record.

The more complex reason has to do with the same nuances of tournament structure discussed in the results sections, and Figure 11. Premier Drafts last significantly longer for winning players (because of the 7 wins or 3 losses structure), which means better than average players play a disproportionately large number of games, mostly against each other. This is what caused the asymmetry in win rates seen in Figure 11. Because of this effect, the modeled win rate (the win rate actually observed based on the changing elo of opponents throughout a draft) for skilled Premier Draft players was closer to the win rate those players actually had when competing against other very skilled players.

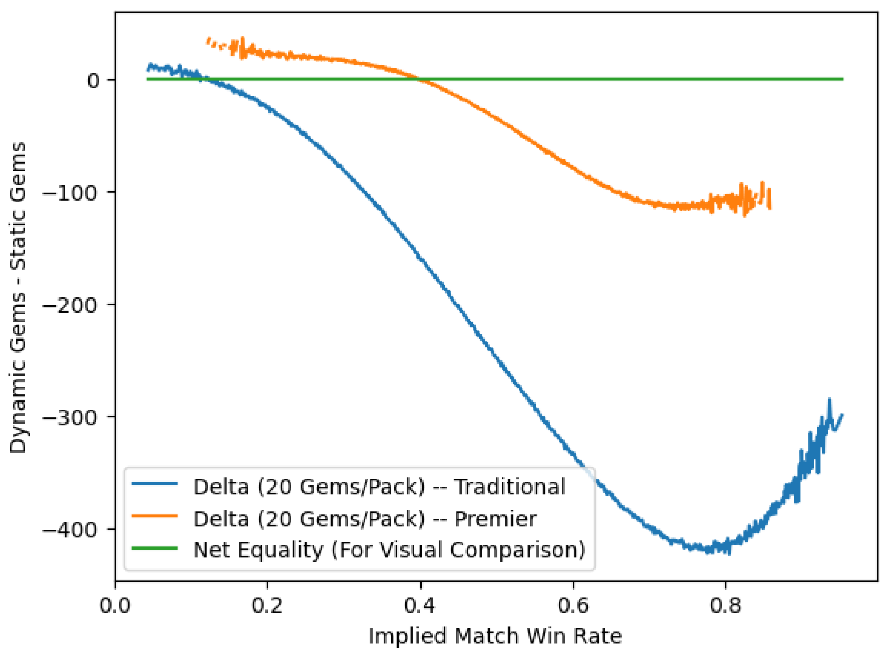
Next, recall that the fundamental problem with the simple, static model is that it un-

[20] I.e, a smaller proportion of the overall prize is distributed to the first place player or best record.

derestimates the extent to which win rate declines as a player performs well in a draft (and vice versa), which causes it to overestimate the likelihood of winning or losing streaks (which we saw in Figure 22). For Premier Drafts though, a disproportionate number of games are played in a skill bracket closer to a player's actual skill level, because each player's average draft lasts a variable amount of time. Combining these effects, the static model overestimates how many trophies a skilled Premier Draft players earns, but it also overestimates how many drafts they lose outright after only winning a couple games. This overestimate of very bad drafts is even worse than simple chance might suggest, because their modeled win rate is reflective of the skill levels of opponents they won't actually face until more easily clearing the first few wins in each draft.

To isolate these effects, Figure 26 below shows the difference in net draft earnings between the static and dynamic model if **implied** match win rate is used to static model prizes, instead of the modeled match win rate observed during drafts (recall that the implied match win rate is the likelihood of winning a match against a 1000-elo opponent).

Figure 26: Difference in Net Cost per Draft by Implied Win Rate



By using the implied win rates, the static model isn't able to take into account the tendency of less skilled players to play against other less skilled players, or vice versa. As we might expect, this causes the average prize to rise for lower ranked players (who would have gained an edge in win rate according to Figure 11), and fall for higher ranked players. The static model is still better at capturing Premier Draft prizing overall, due to the flatter prize structure, but removing the asymmetrically large win rate penalty means a much larger change in implied prizes for high skill Premier Draft players than any other group, reinforcing the effect suggested above.

9 Conclusion

Using a dynamic pairing model and calibrated distributions of player skill, I am able to analyze in detail both the expected results and skill dispersion in iterated draft tournaments with multiple different structures. I find that record-based matchmaking reasonably approximates elo sorting, even on a timescale as short as a 3-round tournament, and that the expected win rate penalties due to this elo sorting are also observed, including the asymmetric penalty applied to more skilled players in Premier Drafts, which involve more games for winning players. Given these win rate penalties, I estimate that only 5% of Premier Draft players and 7% of Traditional Draft players are able to on net profit while drafting, and that the average player will need to spend between $187 to $204 to fully collect a set's rare cards. Finally, I reconstructed a simpler static model and compared my results, concluding that a static model significantly overestimates the frequency of maximum win "trophy" drafts by failing to account for changing opponent skill level. Further, I find that a static model is reasonably accurate when predicting net cost for Premier Drafts (due to quirks in prizing scheme and tournament structure), but significantly underestimates net cost for Traditional Drafts and their more standard tournament structure.

References

[1] 17lands public data. URL https://www.17lands.com/public_datasets.

[2] Elo rating system. URL https://www.chess.com/terms/elo-rating-chess.

[3] Magic: The gathering arena – reward distribution drop rate information. URL https://magic.wizards.com/en/mtgarena/drop-rates.

[4] Jørgen Veisdal. The mathematics of elo ratings, 2019. URL https://www.cantorsparadise.com/the-mathematics-of-elo-ratings-b6bfc9ca1dba.