Exploratory Analysis of Competitive and Economic Interactions in Major League Baseball

Matthew Resnick
Matthew.Resnick@Colorado.edu

Yifan Hong
Yifan.Hong@Colorado.edu

December 14, 2020

Abstract

In this project, we aim to look at certain areas of Major League Baseball from the perspective of directed networks, and ultimately understand the structure and dynamics of these networks. We use network representations of play-level at-bats and inter-team contract trading to determine how they relate to the game in economic, predictive, and physical ways. We employ several methods of network creation and analysis to produce meaningful insights into the essence of the game and related processes.

1 Introduction

When it comes to professional sports in the United States, there is perhaps none more numerically tracked and analyzed than Major League Baseball. There are detailed records of every action on the field, every player's game-related statistics, every contract signed and traded, and even, for recent years, physical characteristics of pitched balls such as exit velocity and angular momentum. Often, this data gets aggregated into "flat" statistics, such as earned run averages (ERA) and on base percentages (OBP). Some more sophisticated statistics have become popular in recent years, such as on-base plus slugging (OPS) and walks plus hits per innings pitched (WHIP), but they remain flat in that they cannot truly be broken down or narrowed in scope without losing important information. For a game in which physical conditions can be more easily controlled for, it makes a great deal of sense to represent interactions inside and outside of individual matches using a more dynamic structure such as networks.

In this project, we pick a few areas of the game to look at using the lens of directed networks and we attempt to glean meaningful insights from these networks by looking at their structure and dynamics over time and at different levels of granularity. We begin by studying play-level at bats by building several different directed, bipartite networks of pitcher/batter interactions over an 11 year period and, using different methods of determining intrinsic hierarchies, generate and study rank embeddings for each. We look at the effect of ranking methodology on predictive power and conclusions about implied hierarchies, and we analyze the idea of players moving up and down in skill over time in the context of these hierarchies. Then, we switch over to looking at directed networks formed from the "meta-game" of season-level player contract trading between teams. We build a directed team-level transaction network among all the MLB teams using trade data and our player hierarchies. We examine the predictive power of our transaction networks on determining team success in the regular season and compare our results to those from using flat stats alone.

2 Background and Related Work

While Major League Baseball is one of the most analytic-friendly sports played a professional level, there is very little published work that frames the game and the league itself in terms of networks. There has been some work on batter-pitcher interactions that mainly revolve around ranking, such as [1] which adapts PageRank to rank players at multiple granularities, and [2] which ranks players using a novel random-walk-based method, and then attempts to generalize the interactions. [3] goes in a somewhat different direction, using an analysis of player social networks to determine the effect of players forming social bonds over time of being on the same team. The existing literature on network analysis in baseball has some serious limitations, which we intend to address. For instance, there is a lack of coherent result analysis, often a narrow scope, and importantly, it is all out of date for a game that has changed significantly in the last 10 years.

For our investigation into intrinsic hierarchies within the batter and pitcher networks, we make use of three ranking methods. The first method is SpringRank [4], which treats directed edges as physical springs and determines a rank embedding by minimizing the energy in the system. We also look at PageRank [5] which produces a rank embedding by representing nodes by their importance as seen by a set of random walkers, and we look at BiRank [6], which smooths information in the graph itself by iteratively optimizing over a normalization function to produce an embedding.

3 Play-Level At-Bats

3.1 Methods

"[Baseball] is a haunted game, in which every player is measured against the ghosts of all who have gone before." - Ken Burns

When it comes to competitive sports, perhaps one of the most discussed avenues of analysis is ranking. Even in casual conversation among sports fans, there is a strong urge towards determining who is the best. But determining any sort of hierarchy among professional competitors, especially a hierarchy in which there is one definitive "best," is ambiguous. Player A is better than player B how? And during what time-frame? And importantly, if you have one number to represent this inequality, how can it account for all of the stochastic processes and specific scenarios which occur during play? So perhaps there is no way to create an ideal hierarchy, but we can at least produce hierarchies which have certain desired properties. For instance, another important aspect of professional sports is prediction, and it is possible to produce hierarchies which are best at predicting the outcomes of events. Predicting edges in a direct, bipartite graph is not a trivial matter, and further we would like to limit the scope of the prediction to groups of players, i.e. batters vs batters and pitchers vs pitchers, rather than the physical batters vs pitchers match-ups that make up the actual at-bats.

It is also important to note that here, edge prediction should not be taken as a metric for *future* prediction, but rather as a metric for two other things: 1) The relative value of a player within their specific group during a given season, and 2) Our confidence in conclusions we make about the hierarchies we discover, especially in regards to skill complexity levels. The first perspective will be especially useful in the later section on evaluating contract trading.

3.1.1 Weighting Events

We start by selecting our domain for prediction: the play-level at-bat (AB). We obtained the publicly available data for every time a batter approached a plate in a major league game between the years 2009 and 2019, and recorded the specific outcomes. Without having done anything but obtained the raw data, this is where we come to the first significant road-block; if we want to know the strength of a result, in addition to the winner of the at-bat, then we have to select a valuation for each result. Certain properties are clear, such as home runs being worth more than walks, but other properties are not, such as how much more a home run should be worth. We came up with two scoring schemes and compare them. The first scheme is to use hand-crafted values for each result based on how we personally want each result to compare. This adds a strong bias to the results, but we hope this bias will be diffused somewhat in the aggregate of all plays. For batter-sided and pitcher-sided results, we provide the following scores (explanation in the Appendix in the repository): **Hit-by-Pitch (1), Walk (2), Single (3), Double (6), Triple (9), Home Run (12), Fielders Choice (1), Field Out (1), Other/Misc. Out (1), Force Out (1), Grounded into Double Play (2), Strikeout (6)**.

The second scheme we used for scoring makes each result's value dependent on its frequency of occurrence during a given year. We still scaled each score by a fixed amount to keep the ordering of events roughly the same, but fluctuation from year to year is now possible. For both scoring schemes, it is unnecessary to consider the other group's scoring because the weights of the edges between groups will not be considered at any point. So we need not be concerned with whether a double is worth as much as a strikeout.

3.1.2 Network Representations

Next, we'd like to compare the efficacy of each scoring scheme, but we first need a method of actually producing a hierarchy. We choose three proven methods for comparison: SpringRank [4], PageRank [5], and BiRank [6]. These were chosen in part due to their success in edge prediction in other domains, and in part due to the fact that they all operate in vastly different ways. But before we can use these, we need more specific networks. The at-bat data along with our scoring methods are fine enough to create a directed bipartite network where we sum up the score for each unique player/opponent pair, but only BiRank can make use of this. So we produce batter and pitcher-only directed unipartite networks by comparing players with the opponents of other players. So if Batter A and Batter B both see Pitcher A, we create a directed edge between Batter A and Batter B weighted by the difference of their scores against Pitcher A. If this weight is positive, then there is a new edge Batter A → Batter B, and if it's negative, the edge is instead Batter B → Batter A. We can then create just one edge by taking a total for each direction, over all pitchers seen by both batters. This scheme produces parallel edges, so to remove this we simply take the difference in weight of the edges. So if our final edges over all common pitchers were A → B (weighted by 10) and B → A (weighted by 5), the edge we will use is A → B (weighted by 10-5=5).

3.1.3 Choices of Scope

Now we are almost able to produce and investigate some actual results, but we need to select the scope of our analysis. We made one choice already, which was to restrict the time-frame to 2009-2019, but we still have to decide at what level we will look at generated hierarchies. There are a lot of meaningful choices we could make in this regard, but to keep the scope of this project focused, we will look at differences in year, pitch type, and inning. Pitch type is an important aspect of an at-bat to control for because both batters and pitchers have their favorite pitches, and it may reveal some difference in hierarchy when we restrict our networks to just one type of pitch. Inning also plays an important role in the results of an at-bat for several reasons, such as batters getting a sense of a certain pitchers style (or vice versa), physical/mental exhaustion, and substitutions, like pinch hitters or relief pitchers, coming into the game fresh, and we see how some of these factors influence our results. Note that we do not consider extra innings in our investigation.

3.2 Results

3.2.1 Ranking Scheme

Let us first look at the results of using each ranking method for predicting the presence and direction of edges in each of the batter and pitcher network over the decade, assuming weights determined by our each of the two scoring schemes. Each ranking method is applied to single seasons with a held-out testing set of 20% of the edges. Then the results are averaged over the entire time period.

Ranking Scheme	Mean AUC (Std. Dev.)	
	Batters	**Pitchers**
SpringRank	0.743 (0.009)	0.816 (0.005)
PageRank	0.370 (0.010)	0.231 (0.008)
BiRank	0.703 (0.009)	0.762 (0.006)

Handmade Scoring

Ranking Scheme	Mean AUC (Std. Dev.)	
	Batters	**Pitchers**
SpringRank	0.750 (0.007)	0.835 (0.005)
PageRank	0.365 (0.010)	0.208 (0.011)
BiRank	0.716 (0.008)	0.775 (0.009)

Frequency Scoring

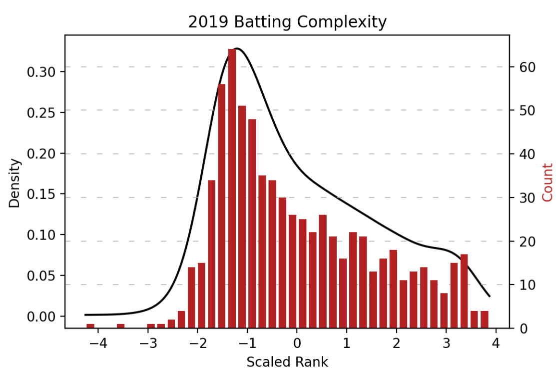
There are several conclusions we can take away from these results, beginning with the relative success of each ranking method at edge prediction. The clearest result is the poor performance of PageRank, though this is not an unexpected result due to the fact that PageRank's operation doesn't make much sense on these networks. The only reason there are directed edges in each of thee graphs is because we produced them arbitrarily from competitive interactions, and thus a random walker has little physical or even theoretical meaning here. SpringRank and BiRank perform comparably well, though SpringRank has the definitive, statistically significant advantage. The quality of AUC for SpringRank is fairly high when considering that the meaning of an edge can be thought of as relative value of a player over an entire season, and thus a 75-80% success rate of determining the value of one player over another. Of course, getting a higher AUC is always desirable, but we discuss in the next section why this may not be possible.

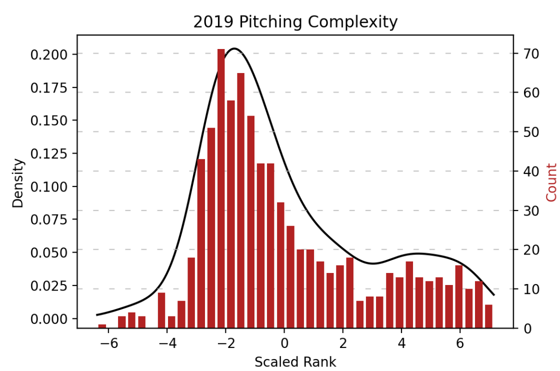
Another conclusion we can take from these results is that there is a slight increase in the AUC for the best performing ranking scheme —SpringRank. However, for the rest of this report we use the handmade scoring scheme to evaluate hierarchy, as the difference in AUC is not enough to justify the increased computational cost during the creation of each scored network. For batters, the mean results are within one standard deviation of each other, and for pitchers there is a larger increase that exceeds three standard deviations, but the overall difference in means is still less than 2%.

Now that we have a solid method of producing rankings for which we are decently confident in, we look at the levels of complexity which arise from it. We fit an inverse temperature to the generated rankings such that one level of skill complexity constitutes the probability that 75% of competitions between those in the level below and those in the level above results in a win for the latter group. We look at the total number of levels which arise for both scoring schemes over the decade:

Scoring Scheme	Mean Number of Levels (Std. Dev.)	
	Batting	**Pitching**
Handmade	8.35 (0.99)	13.86 (0.41)
Frequency	10.96 (2.54)	19.05 (1.06)

These results validate a conventional understanding of the game, which is that there is a higher complexity to the skill of pitching versus the complexity of the skill of batting, in the the number of levels for pitchers is significantly higher than that of batters regardless of scoring. These results also reinforce our choice to use handmade scores for the remainder of our analysis due to the fact that the complexity levels of the frequency scores are much more variable, which makes inter-season comparisons more difficult. We can look directly at the distribution of these levels for handmade scores in 2019:

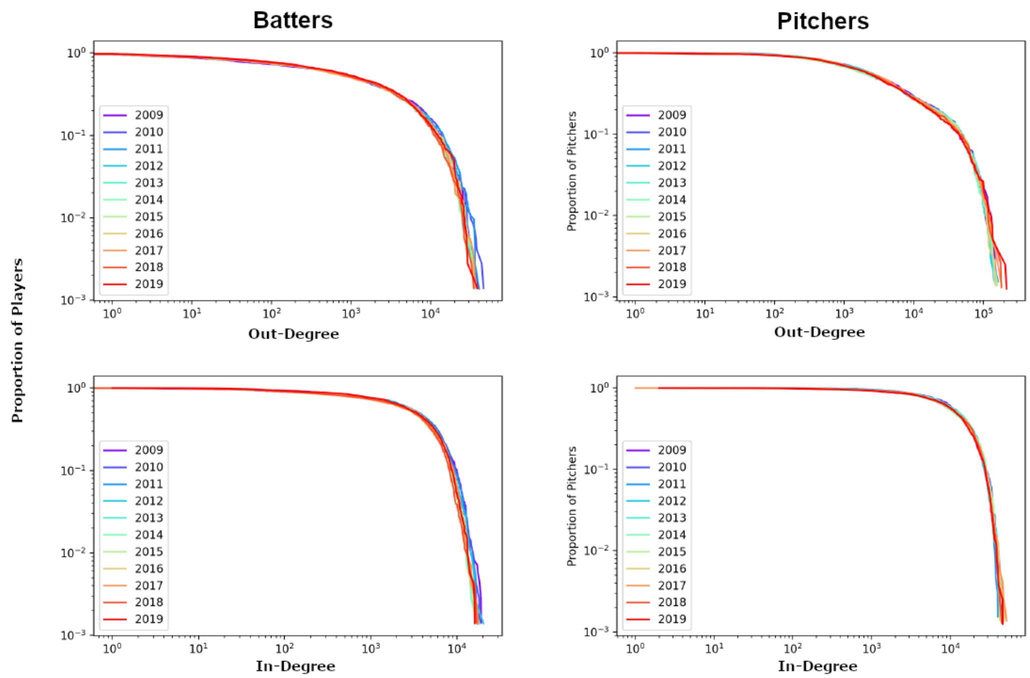




Interestingly, the already narrow batting skill levels have a few outlier players on the low end widening things. It appears that the reason for this is that National League (NL) pitchers were required to bat prior to 2020, whereas hitters essentially never pitch. That's not to say these pitchers can't be good hitters, but they certainly are not signed for their ability to hit. In fact, it seems that much of the low batting tail is composed of NL pitchers. There is a similar low tail for the pitchers, but this appears to simply be due to the sometimes low amount of appearances certain pitchers are able to make in a given season.

3.2.2 Structure

Next we take a brief look at the structure of the networks from which we derived hierarchies in order to add some context to our results. In consideration of degree, we note that batting and pitching over the last 11 years appears to be fair:



(Top) CDF for win strength for batters and pitchers. (Bottom) CDF for loss strength for batters and pitchers.

The weighted in- and out-degree distributions, representing loss and win strength respectively, for both batters and pitchers is not scale-free. Further, it can be seen from the CDF plots that there is generally only a very narrow power-law tail every year, and hence most pitchers and betters tend to be fairly evenly matched in win and loss strengths. This more even-matching in the last decade may, in part, explain the difficulty in predicting edge directions. Predicting players which are distant in embedding to begin with is likely much easier, but as players are closer together, the separation "learned" via the embedding becomes weaker.

3.2.3 Narrowed Domains

What if we restrict our rankings to specific areas of the game? We wish to know if it is easier or harder to make predictions in these regimes, and further how each regime compares to the others. The results via SpringRank can be seen below.

Pitch Type	Mean AUC (Std. Dev.)	
	Batters	**Pitchers**
Changeup	0.652 (0.012)	0.701 (0.014)
Curveball	0.667 (0.020)	0.690 (0.015)
Cutter	0.683 (0.015)	0.702 (0.029)
Four-Seam Fastball	0.670 (0.010)	0.722 (0.007)
Splitter	0.699 (0.072)	0.690 (0.030)
Two-Seam Fastball	0.667 (0.006)	0.697 (0.011)
Sinker	0.680 (0.008)	0.719 (0.012)
Slider	0.648 (0.010)	0.704 (0.007)

Inning	Mean AUC (Std. Dev.)	
	Batters	**Pitchers**
1	0.659 (0.008)	0.698 (0.009)
2	0.644 (0.011)	0.666 (0.012)
3	0.642 (0.007)	0.655 (0.010)
4	0.637 (0.010)	0.656 (0.012)
5	0.638 (0.011)	0.635 (0.014)
6	0.632 (0.012)	0.617 (0.015)
7	0.622 (0.015)	0.607 (0.009)
8	0.632 (0.010)	0.634 (0.010)
9	0.666 (0.009)	0.661 (0.014)

Our results indicate that, overall, all narrowed regimes are more difficult to make predictions on than at the granularity of a season. However, the average quality of predictions between regimes can be incredibly variable. First, we note that five of the eight pitch types have pitcher mean that is more than two standard deviations above the batter mean, and two more that exceed one standard deviation. Thus, pitch type clearly has a much stronger influence on edge prediction for pitchers than for batters, which is not a trivial conclusion but is certainly sensible given the physical context of these predictions. It seemed from the season-level results that pitcher edges were easier to predict overall, but we do not see the same effect in the inning results, so it is more likely that it is pitch type itself influencing the predictions. However, the inning results are also influenced by physical components of the game, as the difference in mean AUC is as much as nearly 4% for batters and 9% for pitchers. The fact that there is a near-linear decline from the first inning on would suggest the increasing influence of difficult-to-predict factors such as physical exhaustion and players getting used to their opponents' play styles. For both batters and pitchers, the mean AUC then goes back up in the 9th inning, potentially due to the higher concentration of relievers by this point for which these influences have less of an effect, though further detailed investigation would be needed to verify this. Another result to note is the highly variable results for both pitch type and inning: 75% of pitch types and 72% of innings for both batters and pitchers have a standard deviation above 1%, compared to just 16-33% at the season level having such a high standard deviation. This suggests that, season to season, some conclusions drawn from predictions on these regimes will not always hold the same.

We see similarly interesting results when considering the skill complexity in these domains. We see a highly variable range of skill levels for pitch types, and we suspect that this is due to the variable usage of each type from season to season and pitcher to pitcher. Consider the Splitter, which has over 17 levels for batting, 8 for pitching, and standard deviations of over 4 and 1.6, respectively, but this pitch was only officially used a few hundred times a season. Compare this to the much more tame statistics for the Four-Seam Fastball, a favorite of many pitchers, and is usually employed over 100,000 times in a season. Interestingly, for the pitch types used the most, the number of levels between batters and pitchers is rather even when compared to the game at-large, suggesting that both batters and pitchers are able to adapt well to every pitch type, without too much room for one group or the other to dominate with a certain type except at the individual level. Of course, this is only at the professional level, so pitchers and batters weak at one pitch type either don't make it this far or, in the case of pitchers, only use other types of pitches.

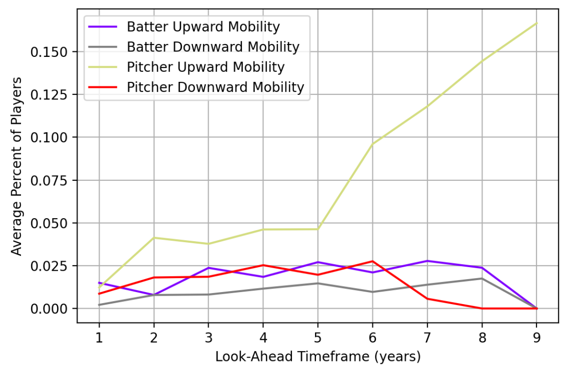
Pitch Type	Mean # of Levels (Std. Dev.)	
	Batters	**Pitchers**
Changeup	6.95 (0.43)	9.13 (1.26)
Curveball	9.28 (1.71)	8.73 (0.68)
Cutter	8.38 (0.76)	8.54 (1.18)
Four-Seam Fastball	7.08 (0.42)	7.73 (0.73)
Splitter	17.47 (4.4)	8.1 (1.67)
Two-Seam Fastball	7.6 (0.9)	8.85 (0.94)
Sinker	7.83 (0.72)	7.91 (1.02)
Slider	6.91 (0.54)	8.42 (1.38)

Inning	Mean # of Levels (Std. Dev.)	
	Batters	**Pitchers**
1	6.61 (0.5)	5.86 (0.77)
2	7.16 (0.54)	6.5 (1.04)
3	7.12 (0.54)	5.85 (0.97)
4	6.98 (0.44)	6.34 (0.64)
5	7.27 (0.58)	5.84 (0.58)
6	7.0 (0.57)	5.34 (0.53)
7	7.29 (0.67)	4.88 (0.52)
8	7.07 (0.75)	6.06 (0.48)
9	8.03 (0.68)	7.71 (0.49)

Since the statistics for inning is much more stable, we can see that there is a slight nuance to the skill levels in different innings. For instance, for batters there appears to be a sizeable difference between the first and last inning, where perhaps more action takes place in the latter in order to avoid a loss or clutch a win. For pitchers there is a similar skill range difference, where it appears there is a wider range of skill for closers than for starting pitchers, either at the beginning of the game or as they begin to lose steam in the middle innings.

3.2.4 Skill Mobility

One question we wanted to answer about the dynamics of skill levels over time is whether players, either batters or pitchers, are able to move up or down in skill level over time. To investigate this, we used our skill level embeddings to look at the average significant movement of players over time from one area of the skill space to another, given some "look-ahead" time-frame. We defined a significant movement as one in which a player moved between the first and fourth quartiles of a given year's skill space. That is, given a look-ahead of two years, if a particular player began in the first quartile of one season's skill levels and ended up in the fourth quartile of the skill levels of a season two years later, we consider this to be upward mobility. We consider downward mobility to be the same amount of movement but in reverse. We then averaged over all players over all years:



Interestingly, it seems that for both groups nearly all skill mobility is quite low regardless of look-ahead frame. Still, we see that looking ahead by 5 years we see roughly 2-2.5% of players move up and down between the top and bottom quartiles. Pitcher upward mobility, however, seems to only become stronger the further we look ahead, indicating that roughly 15% of pitchers in the MLB get better over time. We nearly observe the opposite trend with all other mobility, and this may speak to the deeper skill complexity of pitching over batting and the importance of experience for pitchers.

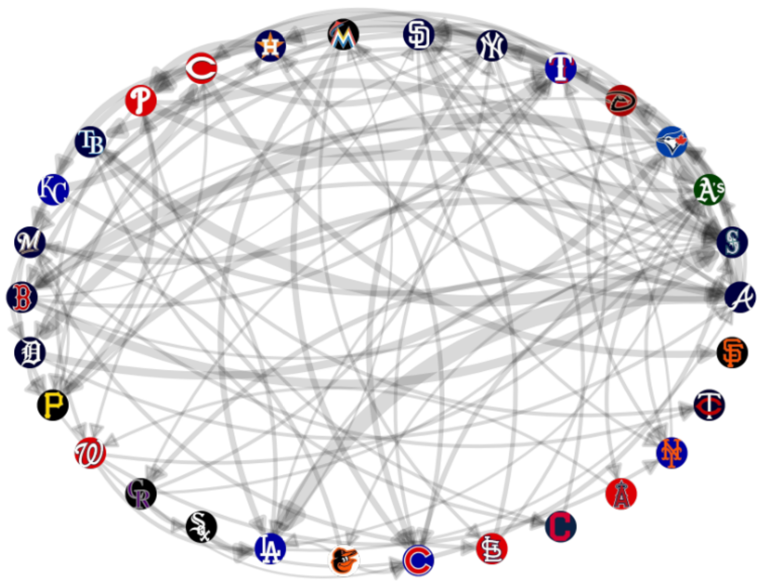
4 Team-level Transactions

4.1 Network representation

All MLB transactions are publicly available, with information detailing the type of transaction (player trade, purchase or draft) and the specific teams/players involved. For the purposes of this project, we examine transactions involving Major League purchases only (discounting Minor League and free agent acquisitions) and record the date, teams involved and name(s) of MLB players. We then construct the team-level transaction network $G_{\text{tr}}(V, E)$ is as follows. Create a vertex for each MLB team (30 total). For each pair of vertices $i, j \in V$, let a directed edge from i to j represent a player transaction from team i to team j. We construct separate networks for every year from 2010-2019 inclusive.

4.1.1 Weighing batter/pitcher transactions

We start by distinguishing each traded player as either a batter or a pitcher during a given year (MLB season). For each identified batter (pitcher), their SpringRank score (handmade scheme) among all batters (pitchers) will be the weight of the transaction. Since the scores in SpringRank are interpreted via relative differences, we shift all the scores by a constant value to remove negative edge weights. The weight of the directed edge from one team to another will then be the sum of the weights of all respective transactions. Thus, our adjacency matrices for annual player transactions are non-negative and we can begin our network analyses.

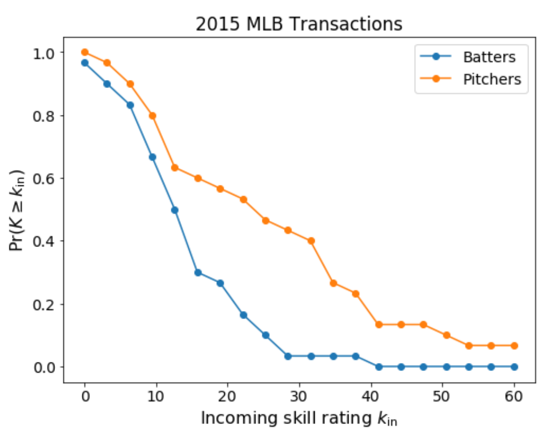


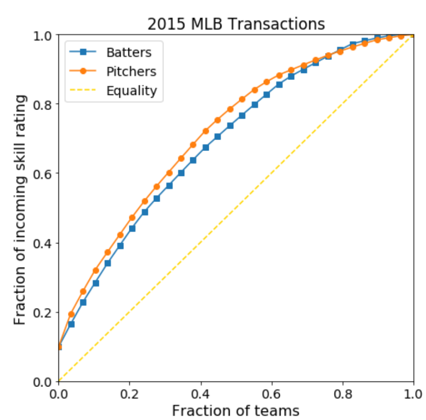
Network visualization for MLB batter transactions in 2015. Each vertex (team) is represented by its respective team logo. Directed edges represent batter skill-rating movement from one team to another, with arrow widths weighted by the SpringRank hierarchy with the handmade scoring scheme.

From initial observations of the network, we can see that teams such as the LA Dodgers and the Boston Red Sox have large in/out degrees, which could be reflective of the large amount of money that these organizations have.

4.2 Network analysis

4.2.1 Cumulative distributions





The left plot shows the cumulative distributions of incoming skill rating in 2015. The right plot contains Lorenz curves of incoming skill rating as a function of MLB teams in 2015.

Edge type	σ_{in}	σ_{out}	G_{in}	G_{out}
Batter Ranks	9.4 (1.5)	9.3 (2.4)	0.354 (0.017)	0.346 (0.025)
Pitcher Ranks	14.9 (3.9)	14.4 (2.6)	0.395 (0.005)	0.380 (0.004)

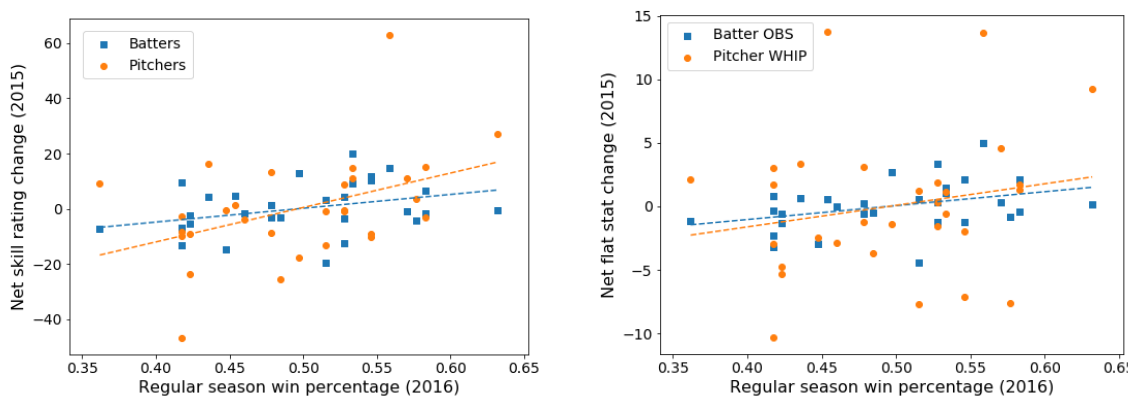
Average in/out degree spread and Gini coefficients are displayed for all MLB teams from 2010-2019.

We observe that the average spread of in/out degrees for team-level pitcher transactions is greater ($\sim \sigma$ significance) than that of team-level batter transactions. This difference in spread can be attributed to the fact that pitching has a deeper skill complexity than batting (from play-level analysis), so we would naturally expect a larger spread in the cumulative distributions for pitcher transactions.

The Gini coefficient $G \in [0, 1]$ characterizes the level of inequality for our degree distributions, with $G = 0$ indicating equality (equal degrees) and $G = 1$ indicating maximal inequality (one vertex contributing everything). For our annual transaction networks, the Gini coefficient represents the inequality in the migration of batting and pitching skill-rating among MLB teams in a given year. We find that incoming and outgoing pitcher transactions are slightly more unequal ($\sim 2\sigma$ and $\sim \sigma$ significance) than incoming and outgoing batter transactions.

4.2.2 Predicting seasonal success

We wish to examine the accuracy of our team-level transactions network at predicting team success during the MLB regular season. We expect a positive correlation between net skill-rating change and win percentage, since we expect teams that gain overall skill-rating to perform better on average than teams that lose overall skill-rating.



Left: Net skill-rating change (Right: Net flat stat change) from 2015 transactions as a function of 2016 regular season win percentage for all MLB teams. Dashed lines represent the lines of best fit to their respectively colored data.

Edge type	Pearson correlation coefficient (Std. Dev.)	
	Same Year	**Next Year**
Batter Ranks	0.39 (0.14)	0.21 (0.12)
Pitcher Ranks	0.37 (0.11)	0.25 (0.13)
Batter OPS	0.31 (0.13)	0.15 (0.12)
Pitcher WHIP	0.14 (0.10)	0.07 (0.10)

Average Pearson correlation coefficient $\langle r \rangle$ between net skill rating change or net flat stat change and regular season win percentage for all MLB teams from 2010-2019.

The Pearson correlation coefficient, often denoted by $r_{xy} \in [-1, 1]$, is a statistic that characterizes the linear correlation between two distributions x, y, with $r = 1, -1$ indicating complete positive, negative linear correlation respectively and $r = 0$ indicating no linear correlation. When we weigh MLB transactions with the SpringRank scores, we observe a moderate positive linear correlation ($\sim 3\sigma$ from $r = 0$) between net skill-rating change and win percentage. This correlation can be attributed to the fact that we are comparing ranks based on play-level data to win percentage of the same season, so a certain player's high skill-rating that year directly contributes to their team's success. We observe a slightly smaller positive linear correlation ($\sim 2\sigma$ from $r = 0$) when we compare net skill-rating change with win percentage of the following year, which represents the predictive power of our player skill-ratings for team seasonal success based on trades and previous-season skill-ratings.

For comparison, we perform the same network analysis as above but with "flat stats" for each player as the weight of each transaction (OPS for batters and WHIP for pitchers). Since higher OPS is better for batters, we would expect some sort of positive correlation with win percentage. A lower WHIP is better for pitchers, so we would expect some sort of negative correlation. We do observe a similar positive linear correlation for batting when we compare net degree with same-year win percentage, but for pitching, we actually see a very small positive correlation on average. The predictive power of these flat stats is also lacking in terms of next-year win percentage (within $\sim \sigma$ of $r = 0$). Thus, by analyzing play-level data to rank players, we obtain edge weights that are more predictive of team success than simply those from flat stats alone.

5 Conclusion and Future Work

In this project, we investigated the structure and dynamics of two important areas of Major League Baseball. We saw that networks derived from game-level at-bats can produce sophisticated valuations of batters and pitchers as well as a space with which we can understand the skill complexity of each portion of the interaction. Further, we were able to make nuanced distinctions of general pitcher-batter interaction dynamics at the granularity of pitch type and inning, and we explored the average movement of players in the space of skill complexity. We then examine the network structure of team-level player transactions, and combine the play-level and team-level analyses to predict team success rate during the regular season. We show that our intricate play-level analysis predicted team success from transactions better than using flat stats alone.

Despite our intriguing results, the space of unexplored areas in the realm of network representations of Major League Baseball is utterly vast, and so there remains much further investigation to be done, both within our selected areas of focus and outside of them. For instance, batter-pitcher interactions are not the only game-level plays which are tracked meticulously, so one further area of research would be to understand the dynamics of fielding plays. Meta-game dynamics also have many avenues to explore, such as incorporating monetary value into trades and expanding recruitment dynamics beyond trades and into minor league, high school, and college baseball.

All code and data found at **https://github.com/mattresnick/MLB-Network-Analysis**.

References

[1] Z. Shan, S. Li, and Y. Dai, "Gamerank: Ranking and analyzing baseball network," in *2012 International Conference on Social Informatics*, pp. 244–251, 2012.

[2] S. Saavedra, S. Powers, T. McCotter, M. A. Porter, and P. J. Mucha, "Mutually-antagonistic interactions in baseball networks," *Physica A: Statistical Mechanics and its Applications*, vol. 389, p. 1131–1141, Mar 2010.

[3] D. Jarvie, "Do long-time team-mates lead to better team performance? a social network analysis of data from major league baseball," *Sports Medicine*, vol. 48, pp. 2659–2669, 2018.

[4] C. D. Bacco, D. B. Larremore, and C. Moore, "A physical model for efficient ranking in networks," 2018.

[5] L. Page, S. Brin, R. Motwani, and T. Winograd, "The pagerank citation ranking: Bringing order to the web," in *Proceedings of the 7th International World Wide Web Conference*, (Brisbane, Australia), pp. 161–172, 1998.

[6] X. He, M. Gao, M.-Y. Kan, and D. Wang, "Birank: Towards ranking on bipartite graphs," 2017.