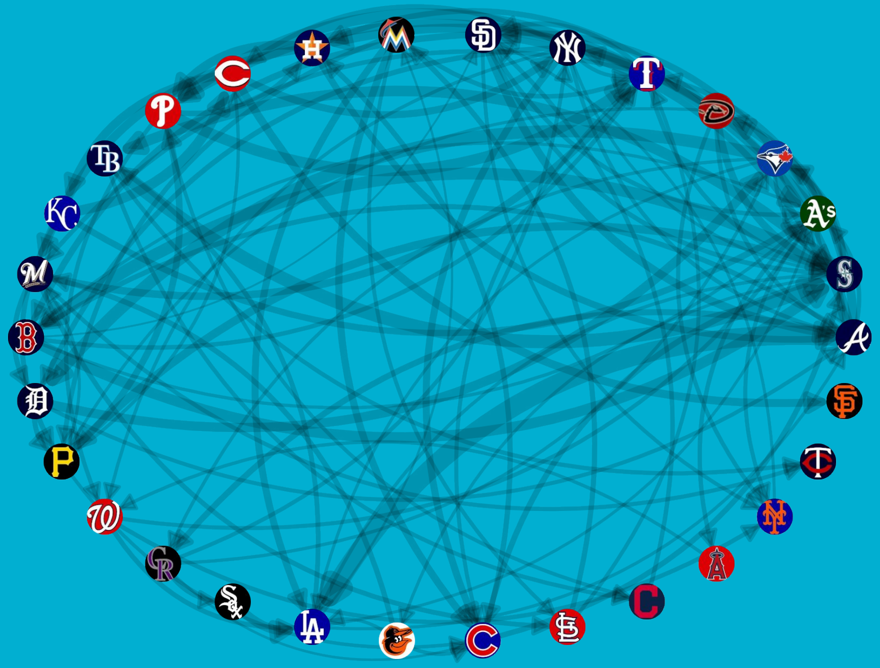
Exploratory Analysis of Competitive and Economic Interactions in Major League Baseball



Matthew Resnick

Yifan Hong

Overview

We produce network representations of two areas of the sport over the last decade to evaluate structure and dynamics:

- Batter-Pitcher interactions in the form of game-level at-bats. Within each group, investigation of hierarchy and skill complexity. ~27 million edges in all related networks.
- Team-level player contract trades weighted by skill-rating hierarchy
  - separate networks for batter and pitcher transactions. ~3000 edges

Related Work

- *GameRank* (Li and Dai, 2012): PageRank to evaluate batter-pitcher interactions and compare with ESPN ranks.
- *Mutually-Antagonistic Interactions in Baseball Networks* (Powers et. Al., 2010): Novel random walk ranking method for batter-pitcher interactions to evaluate sensitivity to rule changes.
- *Do Long-Time Team-mates Lead to Better Team Performance?* (Jarvie, 2018): Social networks of players to predict performance metrics.

We intend to draw more meaningful and interpretable results which are up-to-date with the game.

At-Bat Networks

We begin with discovery of intrinsic hierarchy of at-bat networks.

We build these networks first as directed bipartite graphs (and then ultimately produce single-group unipartite graphs opponent-wise). In the usual way, an edge Batter → Pitcher implies a batter “won” the at-bat.

We then weight the edge by the type of win, and we have two options of doing this: **handcraft** the scores based on a desired outcome, or scale the weights by the **frequency** of the event. We at least want:

- Hit by Pitch < Walk < Single < Double < Triple < Home Run
- Fielder’s Choice ≈ Field Out ≈ Force Out < Ground into DP < Strikeout

At-Bat Networks

Methods of determining hierarchy:

- SpringRank
- PageRank
- BiRank ← Smooths information in the graph via optimization of normalization function.

Choice of scope:

- 2009-2019
- Pitch Type
- Inning (no extras)

Ranking and Scoring Schemes

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

Random walker on our arbitrary networks make little sense, physically or theoretically - PageRank performs poorly.

SpringRank performs best by a statistically significant margin.

Frequency scoring can perform a little better, but not worth the extra computation.

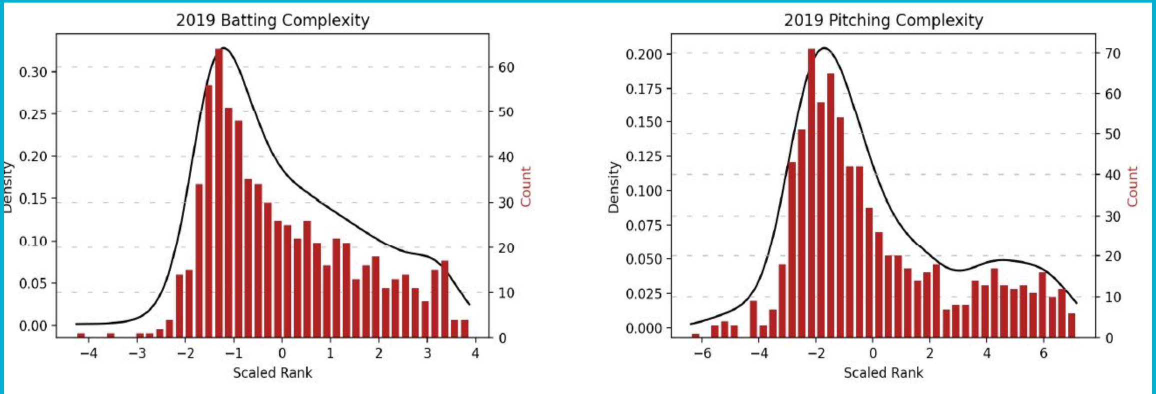
Ranking and Scoring Schemes

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)

Pitching has deeper skill complexity than batting (sensible).

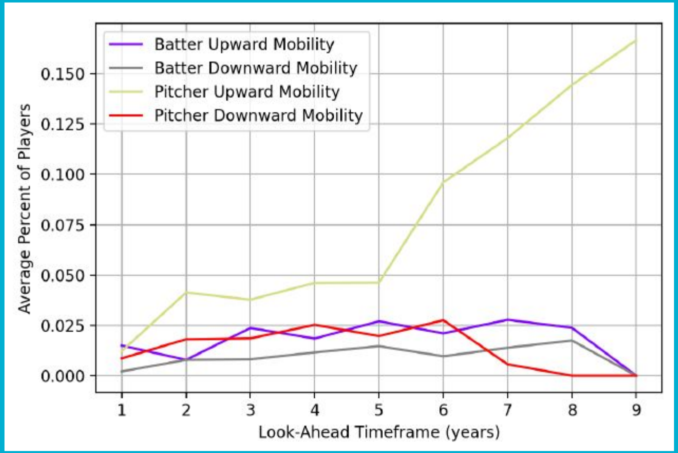
Frequency scoring causes a lot of variation season-to-season (also sensible).

Low tail: NL Pitchers



Low tail: bad showing in few appearances

Skill Mobility



Average movement in skill space between Q1 and Q4.

- Minimal movement in general, max ~2-3%.
- In the far future, batters stay put in skill, and pitchers generally don't go down (or, DFA).
- However, pitcher do move up over longer time periods: more appearances over time + experience matters for pitchers.

Pitch Type and Inning: Edge Prediction

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)

- 5/8 pitch types have pitcher scores exceeding 2σ, 2/8 exceed 1σ. I.e. edge predictions are stronger for pitchers given pitch type.
- Difference in mean AUC is as much as nearly 4% for batters and 9% for pitchers between innings.
- Decline from 1-7 (increasing influence of difficult-to-predict factors such as physical exhaustion and players getting used to their opponents' play styles). Jump at 8 and 9 (relievers?).
- For both pitch type and inning, highly variable results by season.

Pitch Type and Inning: Skill Complexity

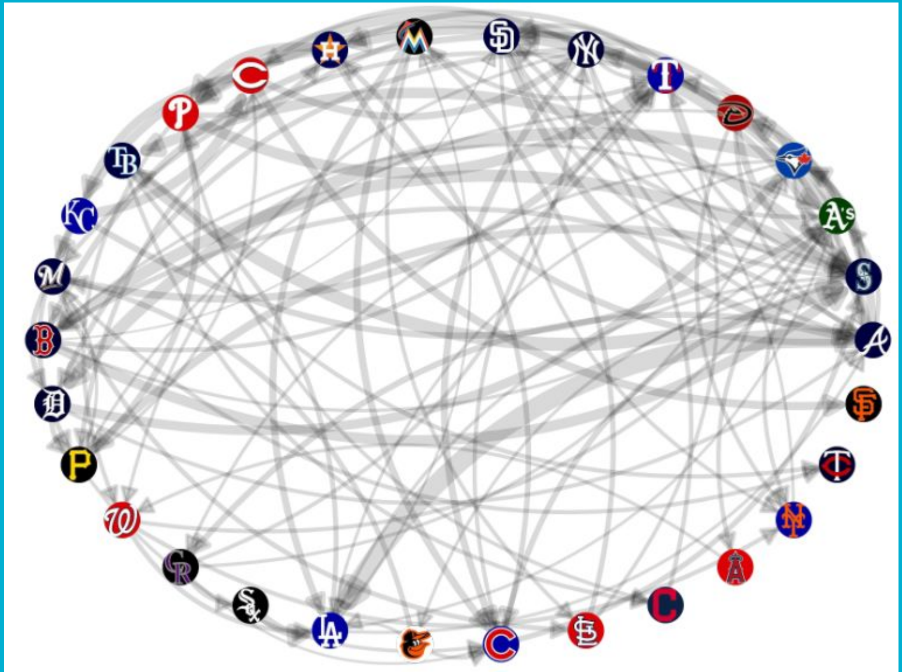
Pitch Type	Mean # of Levels (Std. Dev.)	
	Batters	**Pitchers**
Changeup	6.95 (0.43)	9.13 (1.26)
Curveball	9.28 (1.71)	8.73 (0.68)
Cutter	8.38 (0.76)	8.54 (1.18)
Four-Seam Fastball	7.08 (0.42)	7.73 (0.73)
Splitter	17.47 (4.4)	8.1 (1.67)
Two-Seam Fastball	7.6 (0.9)	8.85 (0.94)
Sinker	7.83 (0.72)	7.91 (1.02)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

- Some highly variable results: FS thrown a few hundred times in a season, FF thrown >100,000.
- Roughly even depth for batters/pitchers across pitch types. Both groups adapt well to the commonly thrown pitches.
- For batters, inning 1 has lower depth than 9. Some can clutch the win, some can’t.
- Pitchers see the same dynamic: greater skill depth to closers than starters.

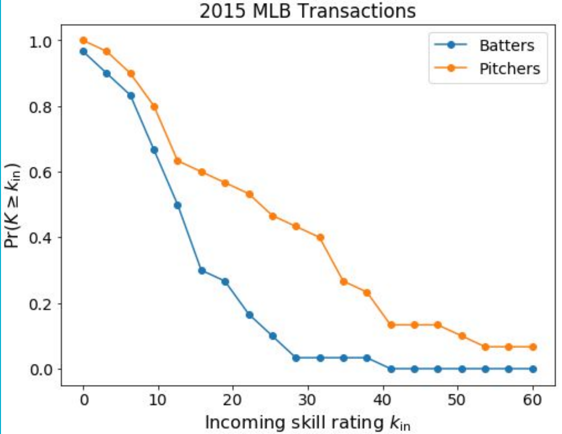
Team-level Transactions

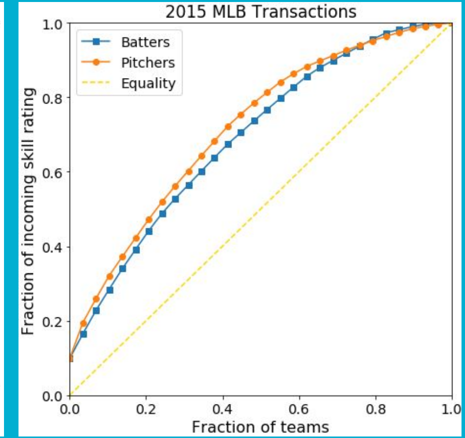
- Publicly available data
- Network representation
  - Player transaction -> Directed edge
- Edge weights
  - Batter hierarchy (SpringRank)
  - Pitcher hierarchy (SpringRank)
- Different network for each year 2010-2019
  - 20 directed networks!



Team-level Network Structure

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)

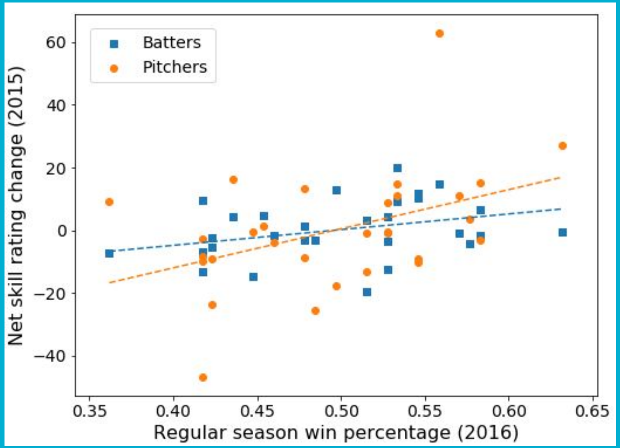




- More degree spread for pitchers
  - Deeper skill complexity
- Inequality among player trades
  - Gini coefficient
  - Pitcher trades slightly more unequal than batter trades

Predicting Seasonal Success

- Does a higher increase in overall skill-rating correspond to a higher win percentage?
  - Pearson correlation coefficient
- Same year vs Next year predictive power
  - Slight positive correlation

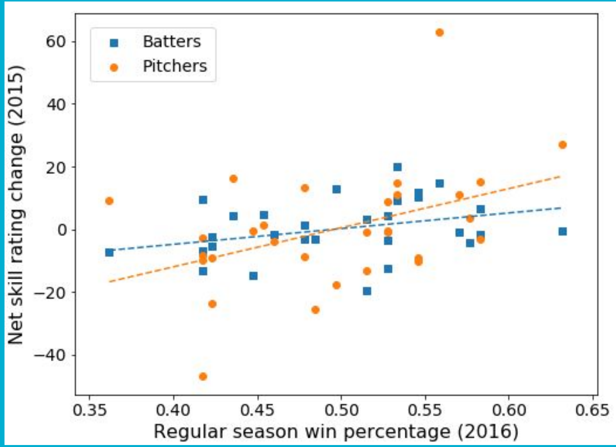


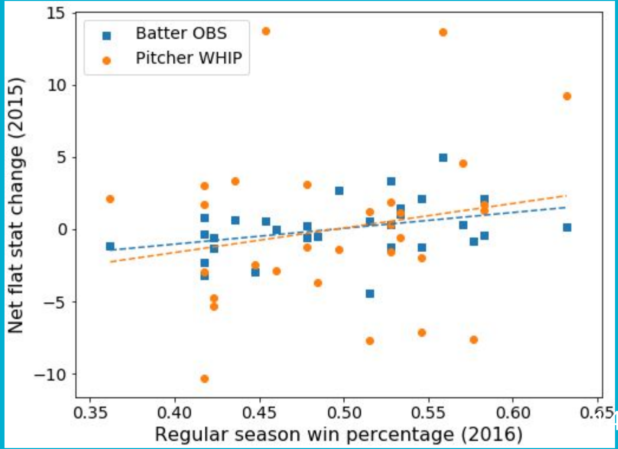
Edge type	Pearson correlation coefficient (Std. Dev.)	
	Same Year	**Next Year**
Batter Ranks	0.39 (0.14)	0.21 (0.12)
Pitcher Ranks	0.37 (0.11)	0.25 (0.13)

Predicting Seasonal Success

- Does a higher increase in overall skill-rating correspond to a higher win percentage?
  - Pearson correlation coefficient
- Same year vs Next year predictive power
  - Slight positive correlation
- Improvement from “flat stats”
  - Batter: on-base percentage + slugging (OPS)
  - Pitcher: walks + hits per inning pitched (WHIP)

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)





Conclusion

Play-level At-bats

- Sophisticated valuations of MLB batters and pitchers from play-level at-bats
- Pitcher-batter dynamics at the resolution of specific pitch types and innings
- Skill mobility over time

Team-level transactions

- Distribution of skill-rating movement among MLB teams
- Predicting win percentage from weighing player contract trades
- Play-level analysis more predictive than flat stats

Outlook

- Interactions outside of batter-pitcher
  - Dynamics of fielding plays
- Incorporating other factors in team-level transactions
  - Monetary value
  - Free agents
  - Recruitment from Minor League and College/High School Draft