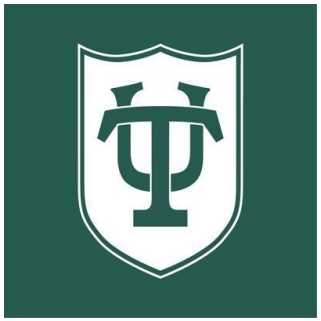
Defining the Most Valuable Player in the NBA

Louis Nass

Tulane University
Department of Mathematics

November 16, 2020



Journalists decide the NBA's Most Valuable Player (MVP)

- Most coveted individual award in the NBA is the most valuable player (MVP)
  - Award recognizes the "best player" for that season
- Journalists vote, player with most votes wins
  - Base voting on **individual player stats**, compelling story-lines, etc.
  - No general definition of "MVP"
- Project goal:
  - Decide which player stats **influence voters** most/least
  - Apply models to 2018-19 season test data

1

Property of the NBA

We must standardize the vote counts before we model

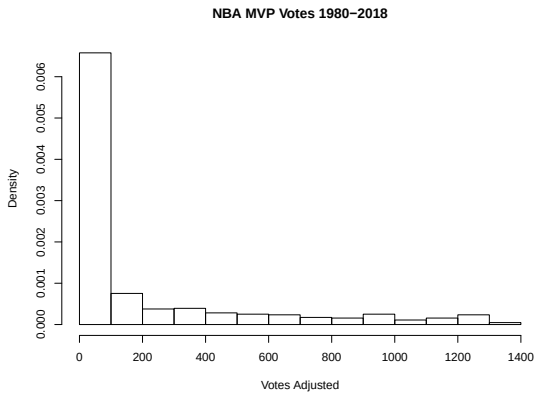
- Data description:
  - Collect NBA individual statistics from 1980-2018 seasons
  - Includes per game stats, games played, PER, etc.
  - Maximum votes received varies each season, **standardize** with *adjusted votes received*

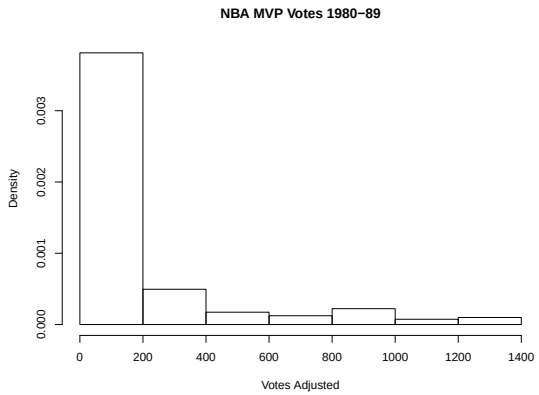
$$\textit{adj. votes received} = \left(\frac{\textit{votes received}}{\textit{max votes avail.}} \right) (\textit{max of max votes avail.})$$

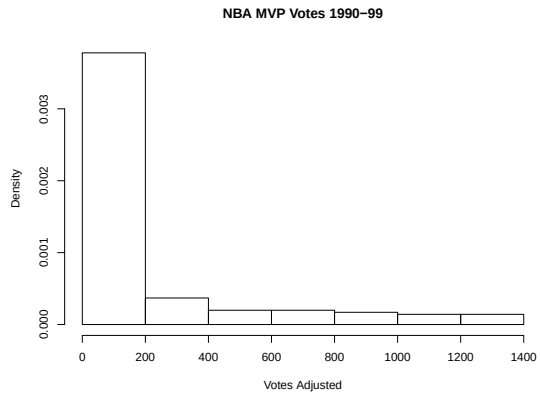
- Project design:
  - Generalized linear models:
    output: *adj. votes received*
    predictors: *individual player stats*
  - 5 total models, one from **all seasons**, one from **each decade** (80's, 90's, 00's, 10's)

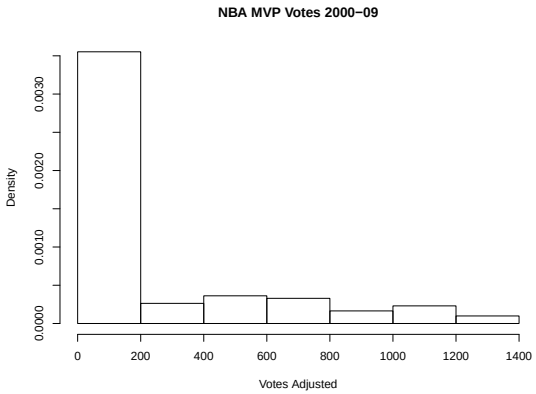
Apply a count regression for the adjusted votes received

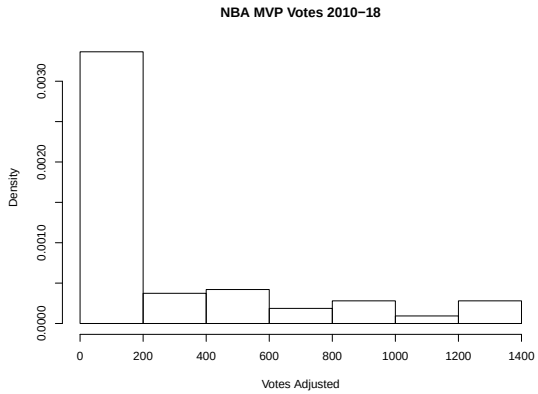
- Since we are counting votes, consider a **count regression**
  - **Poisson** family
  - Votes are **right skewed**, looks like Poisson











Win percent is among the most significant predictors in each model

Top 4 significant player stats from each model

Model	Predictors (ranked 1-4)
1980-2018	Usage percent, **win percent**, games played, minutes
1980-89	**Win percent**, points , field goal attempts, rebounds
1990-99	Usage percent, **win percent**, minutes, win share per 48 games
2000-09	3-point attempts, BPM, **win percent**, assists
2010-18	Minutes, **win percent** games played, win share per 48 games

True shooting is among the least significant predictors

Least significant predictors from AIC criterion

Model	Dropped predictor
1980-2018	**True-shooting percent**
1980-89	Free-throw percent
1990-99	**BPM**
2000-09	No terms dropped
2010-18	**True-shooting percent**, **BPM**, rebounds

These predictors are dropped from our final models

Giannis Antetokounmpo and James Harden are the clear favorites from the models

Vote prediction results

1980-2018		1980-89		1990-99	
Player	Votes	Player	Votes	Player	Votes
Anteto.	7.24	**Anteto.**	7.06	**Harden**	10.03
Harden	7.13	Jokic	5.47	**Anteto.**	7.99
Durant	5.35	Gobert	5.40	Embiid	5.96
Embiid	5.35	Capela	4.97	Curry	5.90
Jokic	5.33	Aldridge	4.17	George	5.80

Giannis Antetokounmpo and James Harden are the clear favorites from the models

Vote prediction results

2000-2009		2010-18		Actual Result	
Player	Votes	Player	Votes	Player	Votes
Harden	8.82	**Harden**	7.18	**Anteto.**	941
Anteto.	7.89	**Anteto.**	6.57	**Harden**	776
Westbrook	6.72	Lillard	5.16	George	356
Embiid	6.26	Gobert	5.13	Jokic	212
Lillard	6.15	Durant	4.98	Curry	175

Harden wins 3 simulations, Antetokounmpo wins 2

How accurate are these results?

Goodness of fit (**Pearson** χ^2)

	1980-2018	1980-89	1990-99	2000-09	2010-18
Pearson χ^2	108084.1	18355.33	16502.78	29589.75	17661.88

- Observe a lack of fit in our models
  - Potential **collinearity** among predictors (i.e. field goal attempts and points per game)
  - Stats from players who **received votes**, not **all players**
- Our models provide **intuition**
  - Projection **mimics** close race between Antetokounmpo and Harden

Figure: James Harden (2018 MVP) Giannis Antetokounmpo (2019, 2020 MVP)

Thank you! Data from Kaggle