

# Optimization of Player Selections in MLB Draft

Jeffrey Ohlmann, Ph.D.



Michael Fry, Ph.D.



Matthew Gibson





# Outline

- **MLB draft**
  - Facts and figures
- **Draft-day decision-making**
  - Factors to consider
- **Simulation/optimization in the draft**
  - Illustration with “Moneyball” 2002 MLB draft
- **Observations**

# MLB Rule 4 Draft



- Every June starting in 1965
- Eligible player pool
  - Residents of U.S., Canada, or U.S. territory
  - First-year “amateur” players
    - Graduating high school players
    - College players eligible after junior year or 21<sup>st</sup> birthday
    - Junior and community college players
- Team retains rights to draftee until August\*
- No draft picks can be traded
  - Picked players cannot be traded for a year after signing

# MLB Rule 4 Draft



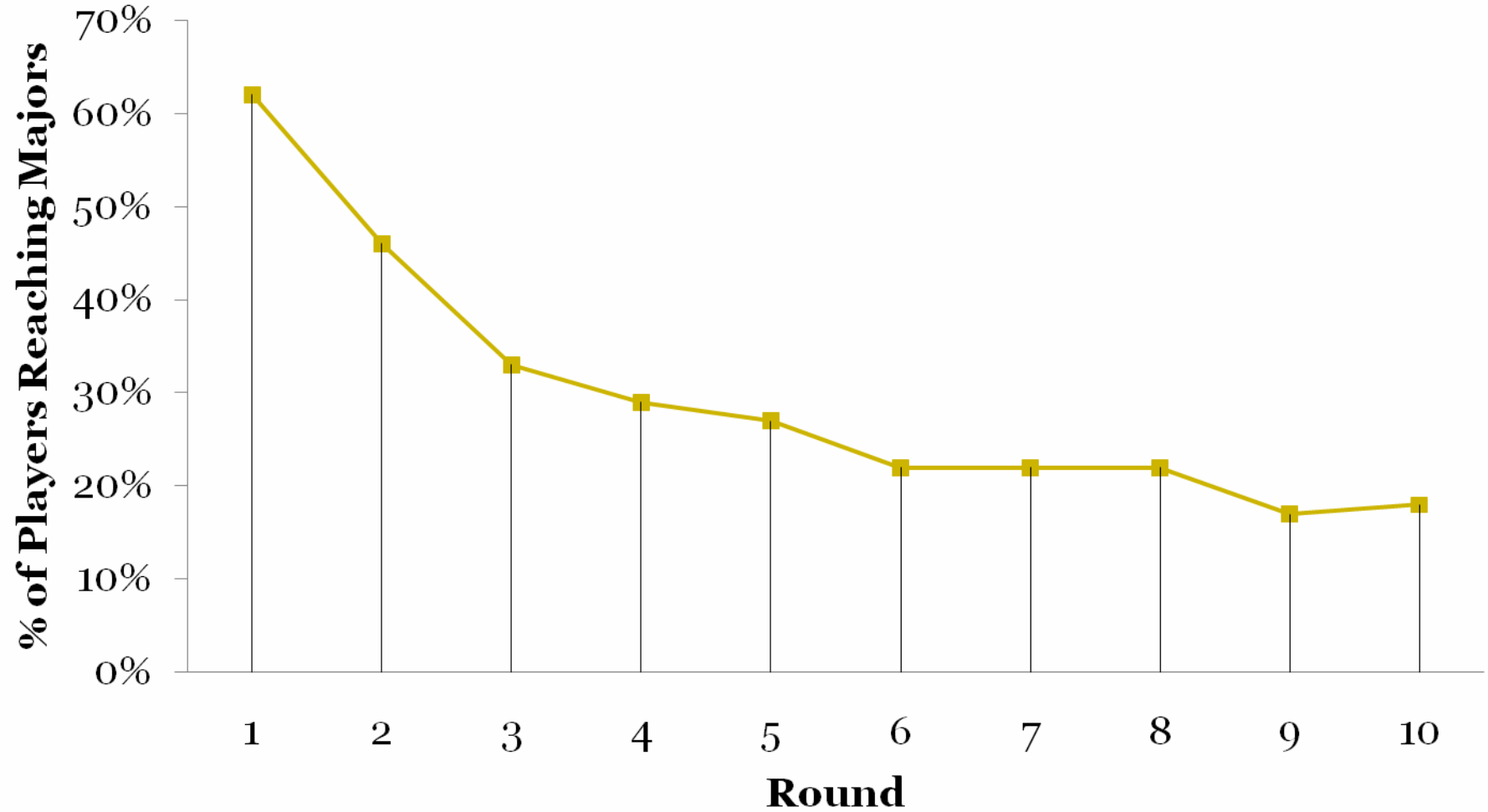
- **Draft order**
  - Based on previous season's standings
  - Compensatory picks for lost free-agents
    - **Type A free agents: top 20%**
      - Supplemental pick between 1<sup>st</sup> and 2<sup>nd</sup> round and compensatory pick from signing team
      - Compensatory pick is highest available pick that signing team has
    - **Type B free agents: next 20%**
      - Supplemental pick

# MLB Draft Summary Stats

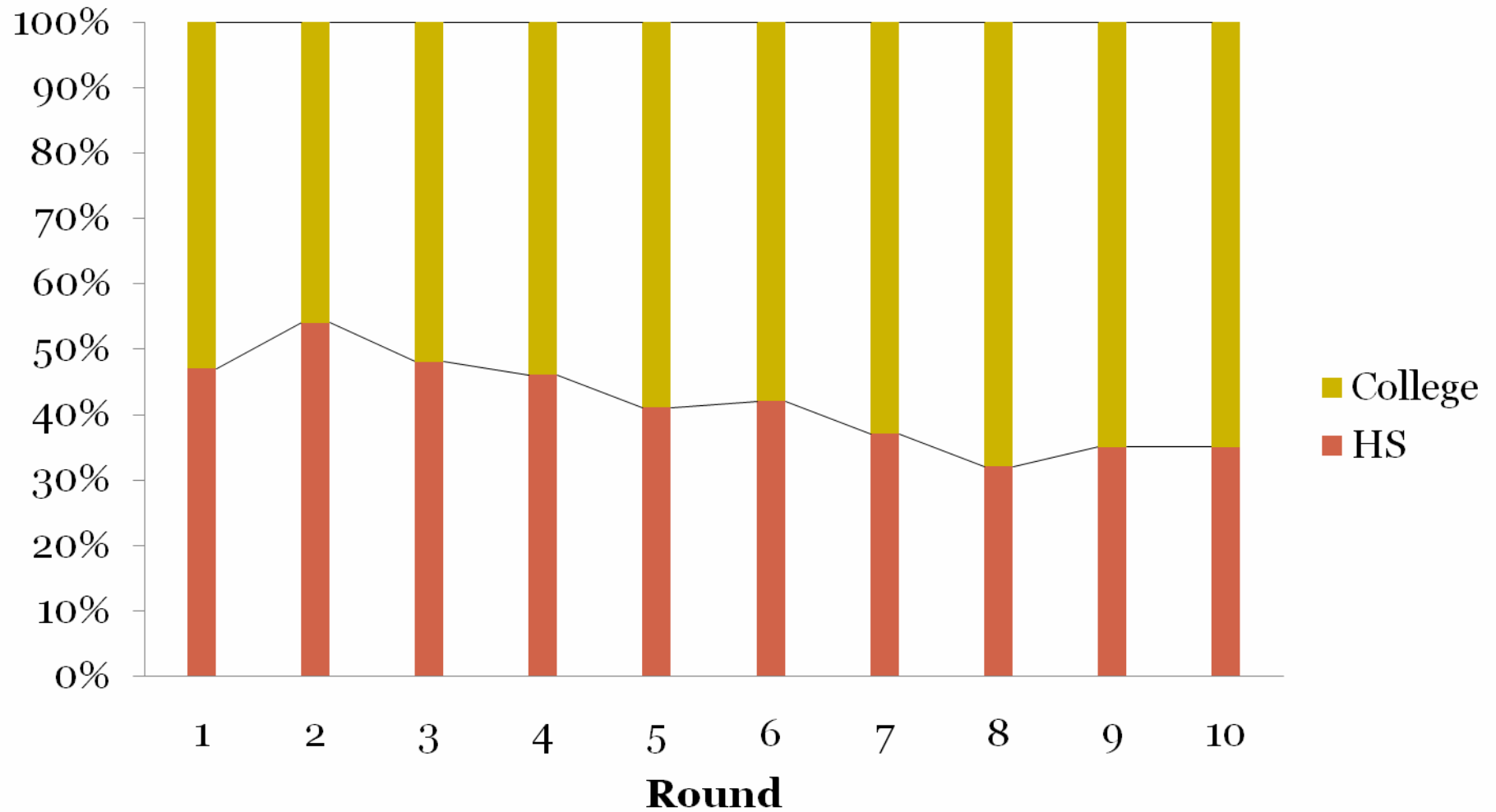


- **Data set**
  - **First 10 rounds of Rule 4 Drafts from 1990-2003**
- **High-level, rough-cut analysis for general characteristics of MLB draft**

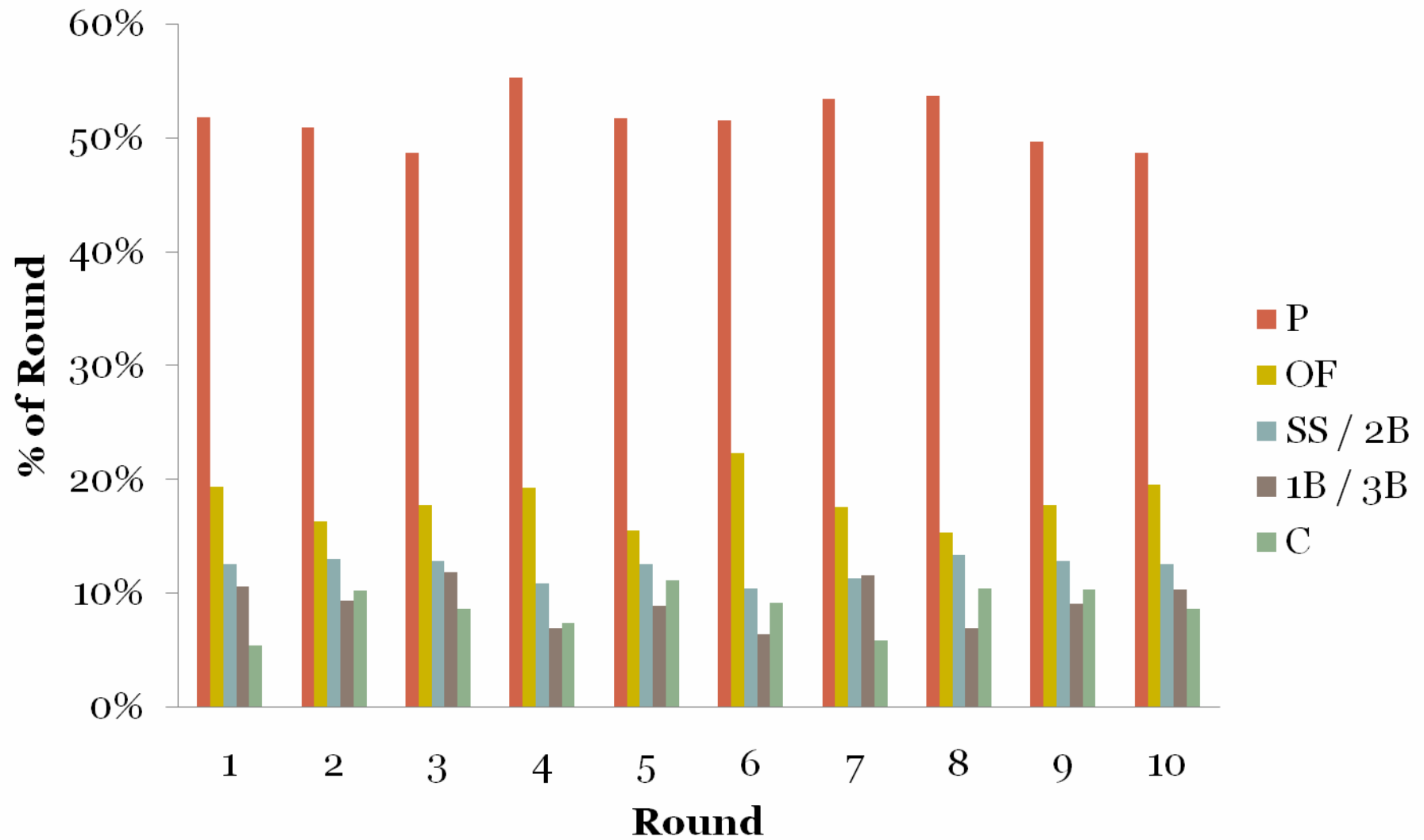
# Drafting Effectiveness



# Drafting High School vs. College



# Positions Drafted



# Draft Success?

Top 5	Inn	Aggregate Top 5	Bottom 5	At Bats / Drafted Hitter
Cubs		Mariners	Mariners	583
Indians		Athletics	Jays	565
Mariners		Twins	Athletics	564
Angels		Rangers	Mariners	562
Athletics		Phillies	Phillies	551

Bottom 5	Innings / Drafted Pitchers	Aggregate Bottom 5	Bottom 5	At Bats/ Drafted Hitter
Marlins		Pirates	Phillies	269
Reds		Reds	Phillies	242
Brewers		Giants	Phillies	222
Yankees		Yankees	Phillies	219
Dodgers		Dodgers	Phillies	195

# Factors on Draft-Day Decisions



- **Decision-maker's player valuations**
- **Organizational needs**
  - Major League roster...who's on the team now
  - Farm system...who's on the way
- **Budget and player signability**
- **How will other teams pick?**
  - Selection strategy of opposing teams
  - Knowledge of other teams' valuations

# DM's Player Valuations



- **How to model?**
  - 1) **Represent as a single number**
    - Average of scouts' grades
    - Maximize the total expected value of draftees' grades
  - 2) **Represent with probability distribution**
    - Mean and standard deviation based on magnitude and consistency of evaluations across scouts
    - Allows for richer modeling, e.g., investment portfolio-type analysis
      - Maximize total expected value s.t. risk constraint

# Organizational Needs



- To what degree should teams draft to fill holes in the farm system or big league roster?
- Model positional needs via cardinality constraints
  - i.e., “need” to draft 5 pitchers in first 10 rounds

# Team Budgets



- **Player signability affect draft selections?**
  - In 2007, Rick Porcello (rated #4 overall by Baseball America) falls to Tigers at #27
  - Evidence not as strong in 2008
    - Royals draft E. Hosmer (Boras client)!
- **Model team budget as a constraint**
  - Drafted players must fit within budget



# Other Teams' Draft Behavior

- How to model other teams' draft selections?
- Knowledge of:
  - Other teams' player valuations
  - Other teams' tendencies and selection strategies (ESPN)
    - Penny-pinchers, safety-seekers, risk-takers, slot-busters
- Use to anticipate and simulate evolution of draft

# Simulating the 2002 MLB Draft



- Use Oakland A's as decision-maker
- Teams' player valuations based on Baseball America rankings
  - Oakland A's player valuations modified to reflect Beane's "wish list" in "Moneyball"
- Team budgets
  - Based on actual expenditures in 2002
  - Player signing bonuses based on max. of slot or actual 2002 signing bonus

# Beane's Wish List

## Pitchers

- ~~J. Guthrie (#10)~~
- J. Blanton (#18)
- ~~J. Francis (#14)~~
- L. Hagerty (#25)
- B. Fritz (#46)
- R. Brownlie (#9)
- S. Obenchain (#170)
- B. Murphy (#70)

## Hitters

- N. Swisher (#35)
- ~~R. Adams (#16)~~
- ~~K. Greene (#26)~~
- J. McCurdy (#49)
- M. Teahen (#134)
- J. Brown
- S. Stanley
- J. Baker
- M. Kiger (#237)
- B. Stavisky (#227)
- S. Larkin
- B. Colamarino

## Actual Picks

1 (16) N. Swisher (#35)  
1(24) J. Blanton (#18)  
1 (26) J. McCurdy (#49)  
1 (30) B. Fritz (#46)  
1 (35) J. Brown (UR)  
1 (37) S. Obenchain (#170)  
1 (39) M. Teahen (#134)  
2 (67) S. Stanley (UR)  
3 (98) B. Murphy (#70)  
4 (128) J. Baker (UR)  
5 (158) M. Kiger (#237)  
6 (188) B. Stavisky (#227)  
7 (218) B. Colamarino (UR)  
8 (248) J. Burton (#129)

## “Optimized” Picks (Clairvoyant)

1 (16) N. Swisher  
1(24) J. Blanton  
**1 (26) L. Hagerty (#25)**  
1 (30) B. Fritz  
**1 (35) J. Votto**  
**1 (37) J. Crain**  
1 (39) M. Teahen  
**2 (67) S. Obenchain**  
3 (98) B. Murphy  
**4 (128) B. Francisco**  
**5 (158) J. Maine**  
**6 (188) M. Kiger**  
**7 (218) B. Stavisky**  
...Brown, Stanley, Baker, Colamarino...

# Observations



- As the DM becomes more uncertain about what other teams will do, the DM will pick more sincerely according to own valuations
- The more picks between consecutive DM selections, the more sincerely the DM will pick
- As the discrepancy between DM's and other teams' player valuations increases, there's more opportunity to take advantage



**Batter Up!**

**Questions?  
Comments?**