

Symposium on Statistics and Operations Research in Baseball

Wednesday, July 16, 2008

<http://ssorb08.sectioninsports.org>

Henry Kaufman Management Center
New York University Leonard N. Stern School of Business
44 West 4th Street
New York, NY 10012

Sponsored by:

Department of Information, Operations, and Management Sciences, Stern School
of Business
American Statistical Association Section on Statistics in Sports
INFORMS Section on Operations Research in Sports

Program

Please note that the program is subject to change.

All activities occur in Room 3-55 unless otherwise noted.

9:15 – 10:00 AM **Registration and Continental breakfast**

Barr-Kawamura Commons, Room 5-50

10:00 – 10:10 AM **Introductory remarks**

John Sexton, President, New York University
Jeffrey Simonoff, Symposium Organizer

10:10 – 11:50 AM **Panel discussion**

Moderator: Gary Hoenig, ESPN Publishing and Editor in Chief,
ESPN The Magazine

Discussants:

Heather Campbell, formerly of ESPN
Michael Kalt, Tampa Bay Rays
Scott Rosner, Wharton Sports Business Initiative
Alan Schwarz, *The New York Times*

Andrew Zimbalist, Smith College

12:00 – 1:30 PM

Lunch break

Barr-Kawamura Commons, Room 5-50

1:40 – 2:10 PM

“Optimization of Player Selections in the MLB Draft”

Jeffrey Ohlmann, University of Iowa

Major League Baseball conducts its First-Year Player Draft, or Rule 4 Draft, annually to assign the rights to amateur baseball players to its franchises. Assuming the perspective of a single franchise, we model selection decisions as an optimization problem subject to uncertainty induced by imperfect knowledge of competing teams' decisions. We discuss the factors which teams consider when making their respective player selections and how they are accounted for in the model. Applying a heuristic search method to this stochastic optimization problem, we determine selection policies and compare them to policies based simple rules-of-thumb.

2:10 – 2:40 PM

“The Value of College: Drafted High School Baseball Players”

Jason Winfree, University of Michigan

This study investigates the decision of high school athletes to pursue college educations or play professional sports immediately after high school. Using salary, draft, and playing data for Major League Baseball, along with minor league playing data, we estimate the lifetime earnings of drafted high school baseball players. The results show that players drafted in higher rounds with more ability should enter professional baseball, but players drafted in lower rounds have higher expected earnings if they attend college. We also find that in 2003, players' decisions were highly correlated with the difference in expected lifetime earnings.

2:40 – 3:10 PM

“Organization of the NCAA Division I Baseball Tournament Championship”

J. Cole Smith, University of Florida

The National Collegiate Athletic Association (NCAA) Baseball Tournament involves 64 teams in a series of win-and-advance weekend tournaments. In the first weekend, a four-team “regional” tournament is played at a host institution. The 16 regionals are also paired with each other (ahead of time, in a bracket), with the winners of paired regionals playing in the second weekend at one of the two institutions' home sites. The eight teams remaining after the second weekend play a final tournament at a neutral location. Given a selection of 64 tournament teams and their seeding classifications, prohibited four-team groupings during the first

weekend, and prohibited regional pairings in the second weekend, we examine the problem of creating regionals and regional pairings in order to minimize expected team travel costs. This talk will examine the performance and output of the proposed algorithm on recent tournament data as compared to the actual tournament given by the selection committee.

3:10 – 3:40 PM

“All-Star Games, Super Bowls, Playoffs and the Regular Season: Effects of Sporting Events on City Coffers”

Dennis Coates, University at Maryland Baltimore County

Cities compete with one another to host important sporting events and even to have teams call them home. This competition is motivated many ways, one of which is the economic benefits to the citizens, the local economy, and to the local tax collector. A wide array of researchers has found little effect of these events on the broad economy, but research into the effects on tax revenues are far less numerous. Using monthly data on sales tax collections from cities and towns across Texas over nearly two decades, there are dramatic differences in the contributions of various sporting events to the revenues of the host cities. My talk will summarize this research and its place in the literature on the economic benefits to communities of hosting sports franchises and special sports events.

3:40 – 3:55 PM

Break

Refreshments available outside room 3-55

3:55 – 4:25 PM

“Forming a Markov Probability Model of 7-Game Playoff Series”

Christopher Rump, Bowling Green State University

We use data clustering algorithms to group similar game-winning frequencies for the favored team in a 7-game playoff series. A Markov chain model that employs these probability estimates provides an extremely good statistical fit to historical playoff outcomes in World Series and League Championship Series play, improving significantly on models based solely on home vs. away game dependence. Surprisingly, even a simple "coin-flipping" model of game-to-game independence does much better than a model incorporating home/away effects.

4:25 – 4:55 PM

“What Do Statistics Tell Us About Steroids?”

Nate Silver, *Baseball Prospectus*

Relative to the amount of attention they have received in the media, the question of Performance Enhancing Drugs (PEDs) has been something of a statistical black box. Were steroids really responsible for the offensive boom of the 1990s? Are their effects on performance all they're cracked up to be? Which types of players are most inclined to use steroids – and stand to benefit the most? Updating and revising the methods from the book *Baseball Between the Numbers*, I tackle these questions and conclude that the effects of steroids on performance are far more ambiguous than is commonly believed.

4:55 – 5:25 PM

“Baseball Cards and the Internet”

Ginger Zhe Jin, University of Maryland

This talk summarizes my research about baseball cards trading on and off the Internet. The study of online trading focuses on the links between eBay price, seller claims of card quality, the actual card quality, and seller reputation. To make a sensible comparison between seller-claimed and real card quality, we purchased actual baseball cards and had them professionally graded. Based on this experimental data and the data downloaded from eBay auctions, we find that some buyers in the online ungraded market are misled by non-credible claims of quality. They pay higher prices but do not receive better quality and in fact are defrauded more often. Online seller reputation is found to be effective for identifying good-faith sellers. But conditional on completed auctions, reputable sellers do not provide better quality. Evidence also suggests that high-claim sellers target less experienced buyers.

To facilitate the comparison between online and offline trading, we purchased the same type of baseball cards in retail outlets and had them professionally graded in the same package. Results suggest systematic sorting of card quality between the online and offline segments. We also demonstrate how retail outlets have changed after the Internet came into place and how supporting industries such as professional grading and card manufacturing adapted to take advantage of the online market.

5:30 – 7:00 PM

Evening reception

Barr-Kawamura Commons, Room 5-50