

Preface

Baseball has always been my favorite sport. My childhood was defined by baseball, from playing Little League, American Legion and in the backyard with my brothers to spending an inordinate amount of time watching the New York Yankees. Growing up in Delhi, New York, I was exposed to Yankee baseball and broadcasters Phil Rizzuto, Frank Messer, and Bill White. Yankee baseball, fishing, and Skoal. With my friend Jeremy Hunter and brother Gerald. Not only did I have to watch every game, but I had to read the next day's NY Post, which I preferred to the NY Daily News.

The teams that I remember were the 1977 and 1978 teams. Mickey Rivers, Willie Randolph, Thurman Munson, Reggie Jackson, Chris Chambliss, Lou Piniella, Graig Nettles, Roy White. I remember cringing often when Bucky Dent came to bat. And of course, Billy Martin. And while my memory is not good enough to remember all of the details, I still remember I was playing ball in the front yard when my brother Michael broke the news of Thurman's death. Perhaps because Munson was my favorite player, I initially despised Reggie Jackson. Only later did I adopt Reggie as my favorite. And I promised to return to Coopers-town when Reggie got inducted; I made it down for the rained out game but not the induction ceremony.

It was in Delhi that I started listening to Pete Franklin's broadcast on 3WE out of Cleveland. When Reggie became a free agent and signed with the Angels, I became an Indians fan. And I have suffered even though I do like cheering for the underdogs. After 1995 and 1997 and the loss of Manny Ramirez, Albert Belle, Victor Martinez, and CC Sabathia, it has been difficult but I still watch the Indians whenever I can. I will admit that I am frustrated with the blackout schedule of Major League baseball. Because I live in Dayton, OH, I am unable to watch the Indians on television because this is the Reds market. The Reds had a good year, making the playoffs, but I would not watch them unless the Indians were in town. I did go see Mark McGwire play in 1998 and enjoyed watching him walk to first base a lot. Thanks, Jack. It makes no sense for MLB to not supply a good when there is a clear demand.

I had the good fortune of working under Jerry Miner at Syracuse University during my graduate days. I was able to look beyond his love for the Yankees; I benefited from his guidance. While at Syracuse, I earned my doctorate degree in economics. I learned the process of research from Jerry Miner and Bill Duncombe. I was fortunate to have the support of my wife Amy, who allowed me to spend an extra year as a graduate student. This allowed me to further develop my research expertise in performance evaluation.

During my interview at the University of Dayton, I met the late Larry Hadley, a sports economist who sparked my interest in applying nonparametric and parametric techniques to analyze sports economics. Larry was a good and generous friend who co-authored several papers with me. Our first paper, co-authored with Elizabeth Gustafson, was on the measurement of technical efficiency in baseball¹ and was published in the 1996 book *Sports Economics*. Elizabeth, Larry, and I followed this up with Gustafson et al. (1999), a paper on econometric specification of baseball production. Larry and I also published a 1999 paper in the *Baseball Research Journal* on evaluating managers. Every time I co-authored a paper we would have the same argument. He insisted that I list my name first while I preferred alphabetical order. He did this because, unlike Larry, I was an untenured junior faculty and he was looking out for my interests. In hindsight, I should have appreciated his generosity rather than find ways to win the argument. But I am happy I won the argument.

Larry and I also teamed up on a few research articles, including Ruggiero et al. (1997), a paper showing that the so-called Pythagoras relationship developed by Bill James cannot be used to measure performance. The relationship between wins in a season can be determined from an identity involving total runs scored, total runs allowed, and the total excess runs. We show the relationship follows algebraically from the simple fact that a game is won if more runs are scored than allowed. Larry and I co-authored a paper with Marc Poitras and Scott Knowles on performance evaluation of NFL teams in 2000. Finally, Hadley and Ruggiero (2006) developed a nonparametric model to evaluate free agents. This model is applied to analyze free agents from the 2009 season.

Unfortunately, I was unable to work further with Larry given his health conditions. I think about Larry often and remember the good times; I will admit that I was somewhat bothered by Larry's ability to choose a favorite team during the playoffs, typically when the Reds were done for the year. Perhaps if I adopted his philosophy I would not be disappointed every year.

My research has benefitted over the years from discussions with many scholars. A special recognition goes to participants at the INFORMS annual meetings who somehow tolerate me. I guess I should extend that to anybody who has ever met me. I have benefitted from conversations with Andy Johnson, Ole Olesen, Timo Kuosmanen, and Tim Anderson. Tim presented a paper at INFORMS to evaluate players; his paper served as a useful reference for my chapter on player evaluation.

¹Ruggiero et al. (1996).

Recently, I have had the pleasure of working on many projects with Andy, who is on pace to become one of the top researchers in performance evaluation. One of our recent papers, co-authored with Trevor Collier, develops a modified performance model useful for sports when inefficiency correlates with other teams' output.² This model is used in this book to measure team efficiency. I also use Collier et al. (2010b) to evaluate individual players.

Organization of the book

The topics in this book are organized into three parts:

- Data envelopment analysis and the evaluation of team performance
- Evaluation of individual players and free agents
- Historical analysis of Hall of Fame selection and the steroid era

The first two chapters of the book are devoted to a brief literature review and the development of the nonparametric models for performance evaluation known as data envelopment analysis (DEA). I do not claim to provide an exhaustive presentation of DEA; this book is intended to be an empirical analysis of Major League Baseball. I present the basic envelopment models to analyze technical and scale efficiency that form the basis for DEA. The literature is full of useful extensions that go beyond the empirical analysis conducted here. In many instances, extensions of the basic models need to be developed for specific chapter applications. Where needed, these extensions are presented in the individual chapters.

The rest of part 1 consists of two applications. I measure the team (and manager) efficiency for the 2009 season in Chap. 3. Cost efficiency for the 2009 season is measured in Chap. 4. Both these models require extensions due to the tournament nature of sports; if a team loses a game due to inefficiency, another team must gain a win. As a result, the estimated frontier from DEA is biased upward. The correction from Collier et al. (2010b) is applied.

The rest of the book is devoted to analyzing player performance. Using a model developed by Collier et al. (2010a), we measure aggregate performance using a modified linear programming model. In the second part of the book, I focus on how DEA can be used to evaluate hitters (Chap. 5), pitchers (Chap. 6), and free agents (Chap. 7). These models provide an overall measure of performance by aggregating multiple player statistics nonparametrically. The methods could be used by teams to make better decisions with respect to draft choices, trades, and free agents.

The last two chapters analyze the performance of players historically. Using DEA, we develop a measure of aggregate performance by season. This allows us to compare how a player performed relative to his peers. Using this measure,

²Collier et al. (2010b).

I analyze the performance of Hall of Fame players and rank the all-time greats. I also identify Hall of Fame players who arguably do not belong in the Hall. I also present arguments on noninducted players who do. The final chapter presents a detailed analysis of steroid use. Using the aggregate measure of performance, I analyze age–performance profiles of numerous players who played before and during the steroid era. Results from admitted steroid users vs. pre-steroid players provide the means to analyze other players. Of course, the results are only suggestive and are only meant for discussion purposes.

Intended Audience

I recently taught a class on DEA as an upper elective for undergraduate and MBA students at the University of Dayton. The material presented throughout this book was at the level of these classes. I spend a lot of time going over the DEA model, linking it to microeconomic production theory while stressing the importance of convexity, monotonicity, and free disposability. The class presents theoretical and methodological extensions that are not presented in this book. Topics covered in this book were used in the class, both as teaching examples and student topics.

The book is accessible to students in economics, mathematics, operations research, industrial engineering, and business programs. The empirical application to Major League Baseball would be useful for practitioners in management, sports management, and sports economics.

Frontiers in Major League Baseball
Nonparametric Analysis of Performance Using Data
Envelopment Analysis

Ruggiero, J.

2011, XV, 139 p., Hardcover

ISBN: 978-1-4419-0830-8