

# Sports Economics, Management and Policy

For other titles published in this series, go to  
[www.springer.com/series/8343](http://www.springer.com/series/8343)



John Ruggiero

# Frontiers in Major League Baseball

Nonparametric Analysis of Performance  
Using Data Envelopment Analysis

 Springer

John Ruggiero  
Edmund B. O'Leary Professor of Economics  
Department of Economics  
University of Dayton  
Dayton, OH 45469-2251  
USA  
ruggiero@notes.udayton.edu

ISBN 978-1-4419-0830-8 e-ISBN 978-1-4419-0831-5  
DOI 10.1007/978-1-4419-0831-5  
Springer New York Dordrecht Heidelberg London

Library of Congress Control Number: 2010937991

© Springer Science+Business Media, LLC 2011

All rights reserved. This work may not be translated or copied in whole or in part without the written permission of the publisher (Springer Science+Business Media, LLC, 233 Spring Street, New York, NY 10013, USA), except for brief excerpts in connection with reviews or scholarly analysis. Use in connection with any form of information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed is forbidden.

The use in this publication of trade names, trademarks, service marks, and similar terms, even if they are not identified as such, is not to be taken as an expression of opinion as to whether or not they are subject to proprietary rights.

Printed on acid-free paper

Springer is part of Springer Science+Business Media ([www.springer.com](http://www.springer.com))

*To my wife Amy  
and our daughters Kailey and Chloé*



# 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 baseball<sup>1</sup> 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.

---

<sup>1</sup>Ruggiero 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.<sup>2</sup> 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,

---

<sup>2</sup>Collier 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.

# Acknowledgments

I would like to first thank the numerous colleagues who have collaborated with me over the years. From Syracuse University, I benefitted from working with Jerry Miner, Jan Ondrich, Johnny Yinger, Bill Duncombe, and Stuart Bretschneider. I owe special thanks to Jerry, my dissertation adviser, without whom I would not be where I am today. Jan was one of my econometrics professors who helped train me and contributed significantly to my development.

I am indebted to Larry Hadley, Ralph Frasca, Elizabeth Gustafson, and John Rapp, all colleagues at the University of Dayton. Ralph was the chair who hired me and provided me with a lot of support. In addition to co-authoring two book chapters, I thank Elizabeth for supporting me as my department chair and interim dean. John has been a quality friend who has been most supportive of me. In the past year I was promoted to Edmund B. O’Leary professor of economics. It goes without saying that this would not have been possible without John’s support. Thank you. I also recognize Dean Matt Shank for his support. In addition to being my current boss and arbitrator over my future raises, Matt is genuinely nice and has a beautiful family.

Another colleague with whom I have written deserve thanks. Marc Poitras, one of the smartest people I have ever met, is a walking encyclopedia of baseball knowledge. I am envious of his ability to recall baseball statistics and other facts.

I have used many sources. I am indebted to Sean Lahman; his extensive database was the primary source for my analysis.<sup>1</sup> I also referenced data from <http://www.mlb.com>, <http://go.espn.com>, and <http://www.baseball-reference.com>. In addition, I used two books by Bill James, *The New Bill James Historical Baseball Abstract* (2001) and *Whatever Happened to the Hall of Fame?* (1995).

Numerous online articles on steroids in baseball were accessed; I have tried to provide footnote references when information was used. In preparation for the steroid chapter, I also read Jose Canseco’s books *Juice* (2005) and *Vindicated* (2008). Both books were enjoyable. In hindsight, it is hard to believe the initial

---

<sup>1</sup>His data can be downloaded from <http://www.baseball1.com>.

negative reaction to *Juice*. It is also hard to believe how many lies have been told regarding steroid use. I benefitted from reading two other books on steroid use. *Game of Shadows* (2006) by Mark Fainaru-Wada and Lance Williams provided useful information on Barry Bonds. Kirk Radomski's *Bases Loaded* (2009) was another interesting book rich in details about steroid use and abuse. These books provided details on steroid use that allowed me to focus on outlier performance.

Over the years, I have benefitted from numerous conversations on baseball. It is impossible to list all, of course, but I will point out a few. My school bus rides with Pete Mansheffer were made enjoyable by Pete's knowledge of baseball. Thanks Pete. I have benefitted from useful discussions with Brent Macintosh, Marc Poitras, Jim and Eileen Malas, Tony Caporale, Jaime Garrett, Gary Abrams, Erica Abrams, Amy Sparks, Randy and Merritt Sparks, Scott Wentz, Jim Gross, and Jacob Fogarty. And I have been enlightened by Bob Thomas, who was a catcher in college. I would also like to thank members of Team John, including but not limited to Amy and Randy Sparks, Sally Doran, Nicole and Scott Wentz, Andrea and Gary Abrams and my wife Amy. Needless to say, their fight for justice in the face of lies is greatly appreciated. Knock it off, Amy Sparks!

I wish to thank my mother Lois, a real baseball fan, and my father Vincent. My mom has given me numerous memorabilia, including her 1949 World Series Program bought at game 1 between the Brooklyn Dodgers and the New York Yankees. Tommy Henrich hit a walk off home run to win the game. My mother was the best little league coach; needless to say we won the championship. For this, I forgive her for still liking the Yankees.

I would like to thank my brother Francis who was my teammate in backyard baseball. He was not gifted with great athletic ability and he did not understand baseball well, but he knew enough to step back to let me shine. I thank him because I suspect he will be the only family member to buy the book.

I also want to thank my facebook friends. Barb and Bob Mackey provided great assistance last year by sending me useful items for my farm. Without their support, I might still be playing farmville instead of writing this book. This has been a team project; all errors are shared by the entire team.

A special thanks goes to Dennis Coates, who asked me to write this book. I met Dennis at the Western Economic Association annual meetings when I presented a paper in Larry's honor. And to my editor Jon Gurstelle, who has consistently supported me, especially by granting me deadline extensions. This has been a team project; all errors are shared by the entire team.

Most of all, I thank my wife Amy and my daughters, Kailey and Chloé. In order to write this book, I have had to put in extra hours that should have been given to them. I promise I will make it up to you by working less in the future.

# Contents

<b>1</b>	<b>Introduction</b> .....	1
	Measuring Technical Efficiency .....	1
	Data Envelopment Analysis .....	1
	Stochastic Frontier .....	2
	Efficiency in Sports Economics .....	3
<b>2</b>	<b>Data Envelopment Analysis</b> .....	7
	Technology.....	7
	Input-Oriented Models.....	8
	Constant Returns to Scale DEA.....	9
	Variable Returns to Scale DEA .....	11
	Output-Oriented Models .....	14
	Measuring Scale Efficiency.....	16
<b>3</b>	<b>Measuring Team Efficiency</b> .....	21
	Serial Correlation .....	21
	Corrected DEA.....	23
	Team Efficiency in 2009.....	24
<b>4</b>	<b>Measuring Cost Efficiency</b> .....	29
	Introduction .....	29
	Cost Efficiency .....	30
	Outlier Correction.....	31
	Cost Efficiency in 2009 .....	33
<b>5</b>	<b>Evaluating Hitters</b> .....	39
	Measuring Aggregate Performance.....	39
	Evaluating 2009 Hitters .....	42
	A Weighted Slack Model .....	44
	Performance and Salary.....	45

<b>6</b>	<b>Evaluating Pitchers</b> .....	49
	Measuring Aggregate Performance.....	49
	Evaluating 2009 Starting Pitchers .....	50
	Evaluating 2009 Relief Pitchers .....	52
<b>7</b>	<b>Arbitration and Free Agency</b> .....	57
	Introduction .....	57
	Nonparametric Estimation of the Contract Zone.....	58
	Hitter's Contract Zone .....	60
	Pitcher's Contract Zone.....	70
<b>8</b>	<b>The Hall of Fame</b> .....	77
	Introduction .....	77
	Evaluating Hitters.....	78
	Evaluating Pitchers.....	84
	Players Deserving Induction .....	86
	Pete Rose .....	89
	Ichiro Suzuki .....	89
	Albert Pujols.....	89
	Joe Jackson .....	90
	Todd Helton .....	90
	Bobby Abreu .....	90
	Barry Bonds .....	90
	Minnie Miñoso .....	91
	Billy Pierce .....	91
	Joe Torre.....	91
	Ron Santo.....	92
<b>9</b>	<b>Steroids in MLB: An Analysis of Hitters</b> .....	93
	Introduction .....	93
	Profiles, Pre-steroids .....	95
	Babe Ruth .....	96
	Hank Aaron.....	98
	Willie Mays .....	99
	Mike Schmidt.....	100
	Pete Rose .....	100
	Profiles, Admitted and Implicated Steroid Users.....	101
	Lenny Dykstra.....	101
	Mark McGwire .....	102
	Jose Canseco .....	104
	Ken Caminiti .....	105
	Jason Giambi .....	105
	Gary Sheffield .....	106
	Barry Bonds .....	107
	Sammy Sosa.....	109
	Manny Ramirez .....	110

- Other Steroid-Era Players..... 111
  - Ken Griffey Jr. .... 111
  - Carlos Delgado ..... 112
  - Jim Thome..... 112
  
- 10 Steroids in MLB: An Analysis of Pitchers..... 115**
  - Introduction ..... 115
  - Profiles, Pre-steroids ..... 116
    - Cy Young..... 117
    - Walter Johnson ..... 117
    - Warren Spahn ..... 117
    - Whitey Ford ..... 118
    - Tom Seaver..... 118
  - Profiles, Admitted and Implicated Steroid Users..... 119
    - Paul Byrd ..... 119
    - Kevin Brown ..... 120
    - Roger Clemens ..... 121
    - Chuck Finley ..... 123
  - Other Steroid-Era Pitchers ..... 123
    - Bret Saberhagen ..... 124
    - Nolan Ryan ..... 124
    - Randy Johnson ..... 125
    - Curt Schilling..... 125
    - Pedro Martinez ..... 126
  
- References..... 129**
  
- Index..... 133**

