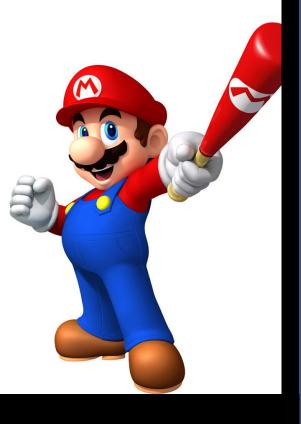


Baseball VI



Produced by Dr. Mario | UNC STOR 538



- Let p = Probability My Team Wins (%)
- Let q = 100-p = Probability Opponent's Team Wins
- How Does an Individual Player Impact p?
- Winning Probability Difference (WINDIFF) $WINDIFF_t = p_t q_t$ where t = Time
- WINDIFF Before Game Begins
 - Assume Each Team Equally Likely to Win

$$WINDIFF_0 = 50 - 50 = 0$$





Jeff Sagarin

- Developed Ranking Methods in Variety of Sports
- Publishes Rankings in USA Today
- Known for MLB Player Win Average Analysis (1957-2006)
- To Avoid Decimals, $SAGWINDIFF_t = \mathbf{10}(p_t q_t) \ \ \text{where} \ \ t = Time$

Scenario

- Home Team Losing by 2 Runs in Bottom of 9th
- Bases Loaded and Mariano Rivera Comes in to Pitch
- Current WINDIFF $SAGWINDIFF_t = 10(52.3 47.7) = 46$
- Batter Hits into Double Play and 1 Run Scores
- Next WINDIFF $SAGWINDIFF_{t+1} = 10(17.2 82.8) = -656$





- Scenario
 - Observe the Change in WINDIFF

$$SAGWINDIFF_{t+1} - SAGWINDIFF_t = -656 - 46 = -702$$

- Clearly, Swing Was in Mariano Rivera's Favor
 - Batter Loses 702 Points
 - Mariano Gains 702 Points
- How Does This Methodology Improve ERA?
- Key Conversion: 2000 SAGWINDIFF = 1 WIN Above 500
 - Changes in SAGWINDIFF For Team = 1000 Points
 - Suppose Team Record is 82W and 80L
 - Across Season

$$82 \times 1000 - 80 \times 1000 = 2000 \ Points$$

Team Ends Season 1 Game Over .500





Win Average Leaders from 2004 to 2006

Sagarin Win Average Leaders, 2004–6					
Year	Player	Position	Total Points	Situations	
2006	Albert Pujols	batter (outfield)	+18,950	653	
2006	Francisco Rodriguez	relief pitcher	+10,562	312	
2005	David Ortiz	outfielder/ designated hitter	+18,145	718	
2005	Roger Clemens	starting pitcher	+12,590	852	
2004	Barry Bonds	outfielder	+25,398	637	
2004	Brad Lidge	relief pitcher	+11,906	382	





- Applied to Fielding Ability
 - Use of Rating System by John Dewan's Fielding Bible
 - Derek Jeter's Rating was -34 (Caused 34 Hits More Than Average)
 - Derek Jeter Cost the Team

$$-34 \ Hits = -0.8 \times 34 \ Runs = -27.2 \ Runs$$

$$= -\frac{27.2}{10} \ Wins = -2000 \times 2.72 \ Win \ Points$$

$$= -5,440 \ Win \ Points$$

- Adjust Derek Jeter's Win Points by Subtraction
- Historically, Fielding Has Been Overrated
- Based on Fielding Bible, Only 7 Players with Positive Fielding Ratings Equivalent to 2 More Wins Above Average.





- Applied to Baserunning Ability
 - Good Base Runner Described
 - Rarely Caught Stealing
 - Rarely Caught in Double Plays
 - Able to Take Extra Bases
 - Player Win Averages Reward Stolen Bases and Preventing Double Plays
 - Fast Base Running is not Rewarded by Player Win Averages
 - Analyzed By Dan Fox (Director Of Baseball Informatics for Pirates)
 - Compare Runner's Number of Runs to Average Runner
 - Best Base Running Has Little Effect on SAGWIN





- Interesting Application: Hitting Versus Pitching
 - 1969 World Champion New York Mets
 - Table of Player Win Averages
 - Hitters = Bad
 - Pitchers = Good
 - Season 100 Wins
 - Expect 38,000 Win Points

$$(100 - 81) \times 2000 = 38,000$$

Not Equal Because Trades

	Hitters total		Pitchers total	Team total
	-9319		41057	31738
Batter	Winpoints	Pitcher	Winpoints	
Grote	-1960	Seaver	13471	
Kranepool	-765	Koosman	13218	
Boswell	-42	Cardwell	761	
Garrett	-3819	McAndrew	1332	
Harrelson	-131	Ryan	23	
Agee	5410	McGraw	10902	
Jones	6334	Koonce	185	
Swoboda	6278	DiLauro	890	
Shamsky	4998	Taylor	1625	
Weis	-2054	Frisella	-605	
Gaspar	-1772	Jackson	-745	
Pfeil	-3629			
Clendenon	240			
Martin	-2196			
Charles	-1021			
Otis	-1789			
Dyer	-290			
Dyci	270			





- Winning Probabilities for All Game Scenarios
 - Recall: p = Probability My Team Wins (%)
 - How is this Probability Calculated?
 - We Can Use a Massive Set of Play-by-Play Data (1977-2006)
 - Scenarios Based on Inning, Score, Runner Locations, and Outs
 - Problem Some Scenarios Are Rare
 - Advised to Use Markov Chain Monte Carlo (MCMC)





Keith Woolner

- Works for the Cleveland Indians
- Authored for Baseball Prospectus
- Respected for Player Analysis and Market Evaluation
- Created Value of Replacement Player (VORP)

Replacement Player

- Players Create Value by Keeping Bad Players Off the Field
- Acquiring Replacement Players (Minor League)
- Assume List of Replacement Players is Infinite
- Woolner Defined Replacement Player as a Player in the Bottom
 20% of List Ordered by Plate Appearances
- Team of Replacement Players Would Get Approximately 44 Wins and 118 Losses



 Value Over a Replacement Player Points (VORPP) PA = Plate
Appearance
BFP = Batters Faced

- Recall: Each Loss Below .500 = -2,000 SAGWIN
- For All Replacement Players = -74,000 SAGWIN
- Batters and Pitchers Get Equal Blame = -37,000
 SAGWIN Points Each
- Approximately 6,200 Plate Appearances in a Season

$$\frac{-37,000}{6,200} = -5.97 Points Per Plate Appearance$$

Formula for Hitters

$$VORPP = SAGWIN + 5.97 \times (PA)$$

Formula for Pitchers $VORPP = SAGWIN + 5.97 \times (BFP)$



VORPP Used to Evaluate Trades

Hoffman VORPP =
$$7,963 + 5.97(255) = 9,485$$
.
Young VORPP = $6,117 + 5.97(781) = 10,780$.

VORPP Used to Determine Salary

Pujols 2006 value =
$$22,848 \times 1,040 = $23.8$$
 million.
Hoffman 2006 value = $9,485 \times 1,040 = 9.9 million.
Young 2006 value = $10,780 \times 1,040 = 11.1 million.

$$$77M = 74,000 VORPP$$

 $$1,040 = 1 VORPP$





- Was Alex Rodriguez Overpaid by Yankees? Yes
 - Deal: \$275M for 10 Years

Year	Plate Appearances	SAGWIN Points	VORP Points	Fair Salary in 2006 Dollars (millions)	
2003	741	10593	15016.77	\$	15.62
2004	734	6575	10956.98	\$	11.40
2005	752	12521	17010.44	\$	17.69
2006	695	2205	6354.15	\$	6.61
2007	736	14193	18586.92	\$	19.33

Fair Deal Based Off 2007 Statistics

 $$19.33M \times 1.05 \times 10 \approx $202M$

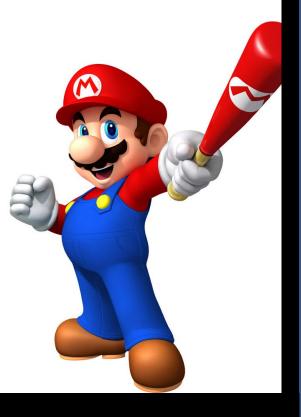




Extra Plate Appearances Create Value

Inge VORPP =
$$15 + 5.97(617) = 3,699$$
.
Therriot VORPP = $2,480 + 5.97(174) = 3,519$.

VORPP Used by Theo Epstein in 2004 Red Sox Season





- Wins Above Replacement (WAR)
 - Attempt to Measure All of a Baseball's Player's Contributions
 - What is it good for? Absolutely Everything
 - Not Comparing to Average Player but Replacement Player
 - Interpretation: "If a player got injured and their team had to replace them, how many wins would the team be losing?
- Three Main Sources of WAR
 - Baseball-Reference (bWAR)
 - FanGraphs (fWAR)
 - Baseball Prospectus (WARP)





- Formula Differs for Position Players and Pitchers
 - Position Players: Batting, Baserunning, and Fielding Measured in Runs Above Average (RAA)
 - Pitchers: Either Based on Runs Allowed Per 9 Innings (RA9) or Fielding Independent Pitching (FIP)
- General Formula for Position Players (fWAR)

 $WAR = (Batting\ Runs + Base\ Running\ Runs + Fielding\ Runs + Positional)$

Adjustment + League Adjustment +Replacement Runs) / (Runs Per Win)





Top Players for bWAR Based on Career

Rank	Player (yrs, age)	WAR Position Players	Bats
1.	Barry Bonds (22)	162.8	L
2.	Babe Ruth+ (22)	162.1	L
3.	Willie Mays+ (22)	156.4	R
4.	<u>Ty Cobb+</u> (24)	151.0	L
5.	Hank Aaron+ (23)	143.0	R
6.	<u>Tris Speaker+</u> (22)	134.1	L
7.	Honus Wagner+ (21)	130.8	R
8.	Stan Musial+ (22)	128.2	L
9.	Rogers Hornsby+ (23)	127.0	R
10.	Eddie Collins+ (25)	124.0	L





Top Players for fWAR Based on 2019 Season

#	Name	Team	Pos	PA	IP	Primary WAR	Total WAR
1	Mike Trout	Angels	CF	600		8.6	8.6
2	Christian Yelich	Brewers	RF	580		7.7	7.7
3	Alex Bregman	Astros	3B/SS	643		7.4	7.4
4	Cody Bellinger	Dodgers	RF	621		7.2	7.2
5	Anthony Rendon	Nationals	3B	595		7.2	7.2
6	Ketel Marte	Diamondbacks	2B/CF	626		7.0	7.0
7	Jacob deGrom	Mets	Р		190.0	6.2	6.9
8	Marcus Semien	Athletics	SS	698		6.8	6.8
9	Max Scherzer	Nationals	Р		159.2	6.4	6.4
10	Gerrit Cole	Astros	Р		192.1	6.5	6.4



Final Inspiration

Yesterday's home runs don't win today's games.

-Babe Ruth