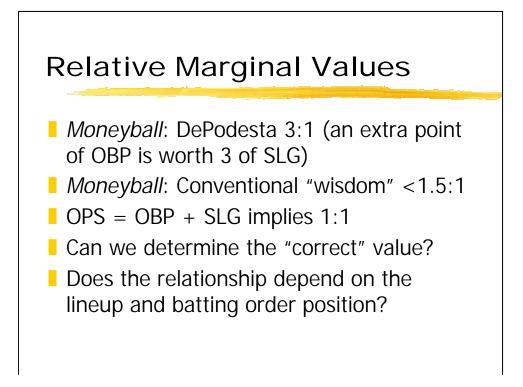# Relative Value of On-Base Pct. and Slugging Avg.

Mark Pankin
SABR 34
July 16, 2004
Cincinnati, OH

Notes provide additional information and were reminders during the presentation. They are not supposed to be anything close to a complete text of the presentation or thorough discussion of the subject.

Relative Marginal Values

- *Moneyball*: DePodesta 3:1 (an extra point of OBP is worth 3 of SLG)
- *Moneyball*: Conventional "wisdom" <1.5:1
- OPS = OBP + SLG implies 1:1
- Can we determine the "correct" value?
- Does the relationship depend on the lineup and batting order position?

Research inspired by comments on page 128 of the book. It says DePodesta tinkered with Runs Created formula to come up with the most accurate runs estimator he knew, and it led to a 3:1 ratio.

Analysis assumes (I think) all batters are the same. In a real lineup, the relationship may well vary by lineup position and the preceding or following batters.
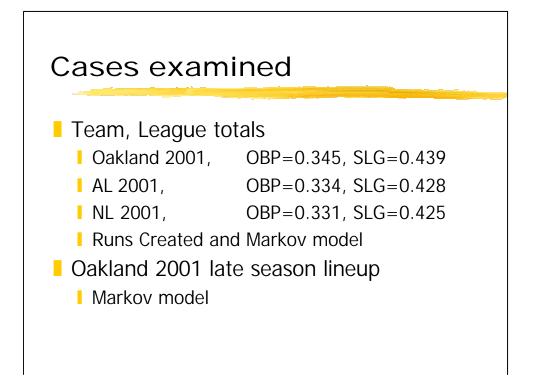
Analytical Approach

- Changing number of BB (for player, team, league) affects OBP, leaves SLG the same
- Changing distribution of 1B, 2B, 3B, HR affects SLG, leaves OBP unchanged
- Change OBP by a specified amount, find change in runs (Runs Created, Markov)
- What SLG change produces same runs?

While not likely realistic when comparing two players or teams, it is possible to vary OBP without affecting SLG by varying the number of walks and to vary SLG without affecting OBP by changing the distribution of hits while leaving the number of hits the same. That is change some singles to extra base hits (in proportion to the actual distribution) or vice-versa.
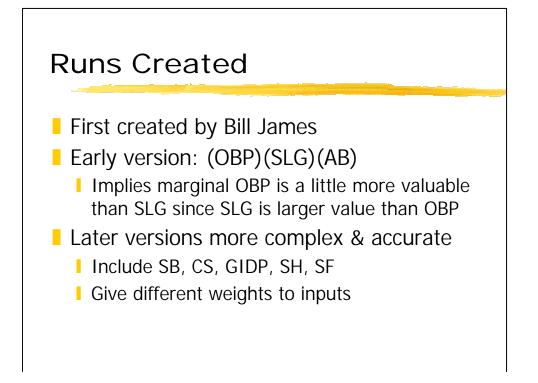
Will change OBP by +/- 10, 20 points, see the effect on runs using RC and the Markov model. Then play with the extra base hit distribution to get the same change (after restoring BB to the original value) and see what the corresponding SLG is.

Note: all information used is available from www.retrosheet.org (disclosure: I am the webmaster)

Cases examined

- Team, League totals
  - Oakland 2001,   OBP=0.345, SLG=0.439
  - AL 2001,   OBP=0.334, SLG=0.428
  - NL 2001,   OBP=0.331, SLG=0.425
  - Runs Created and Markov model
- Oakland 2001 late season lineup
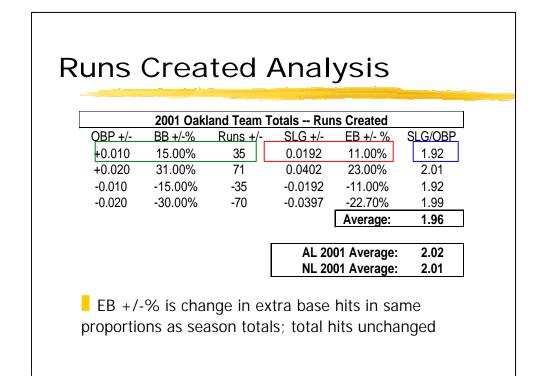 - Markov model

Chose 2001 Oakland due to *Moneyball* focus on that team after the season was done. Wanted to compare to league-wide for that year. Could use other teams and other years for a more complete analysis. Both RC and Markov apply to case when all hitters are the same (team, league average).

Markov can also deal with real lineups, when all the batters are different.

Runs Created

- First created by Bill James
- Early version: (OBP)(SLG)(AB)
  - Implies marginal OBP is a little more valuable than SLG since SLG is larger value than OBP
- Later versions more complex & accurate
  - Include SB, CS, GIDP, SH, SF
  - Give different weights to inputs

Not sure if I used the most recent version, but I did use one or the later ones for the analysis. Ratio of marginal values based on early version is about 1.3.

Runs Created Analysis

2001 Oakland Team Totals -- Runs Created					
OBP +/-	BB +/-%	Runs +/-	SLG +/-	EB +/- %	SLG/OBP
+0.010	15.00%	35	0.0192	11.00%	1.92
+0.020	31.00%	71	0.0402	23.00%	2.01
-0.010	-15.00%	-35	-0.0192	-11.00%	1.92
-0.020	-30.00%	-70	-0.0397	-22.70%	1.99
				Average:	1.96

AL 2001 Average:	2.02
NL 2001 Average:	2.01

▮ EB +/-% is change in extra base hits in same proportions as season totals; total hits unchanged

Walk through first row explaining meaning:
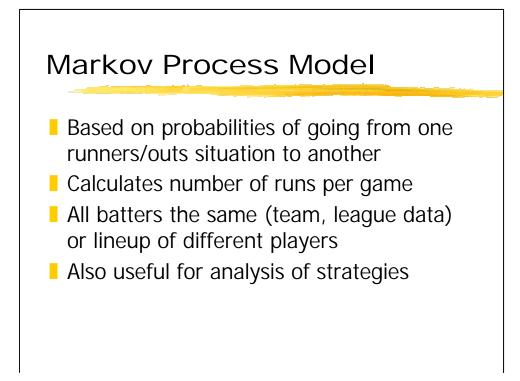
+15% BB to raise Oakland OBP by 10 points

leads to RC model value of 35 runs increase (for season)

by trial and error determine an 11% increase in the proportion of extra base hits has the same effect

that results in a 0.0192 increase in SLG, so ratio of marginal values is .0192/.010 = 1.92. That is an extra OBP point is worth 1.92 extra SLG points.

Note values are not quite "linear".

Four case details not shown for the leagues, just the average of the four cases for each league.

Markov Process Model

- Based on probabilities of going from one runners/outs situation to another
- Calculates number of runs per game
- All batters the same (team, league data) or lineup of different players
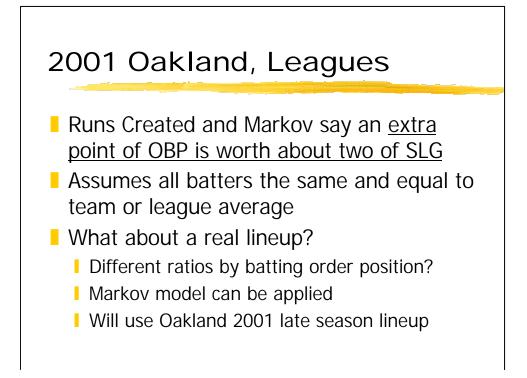- Also useful for analysis of strategies

I have used the Markov model extensively for baseball strategy analysis and have given several talks on the subject at prior SABR meetings. (Last year to see when it makes sense to walk Bonds.) It is well suited to study the OBP vs. SLG question.

The model version used incorporates ML averages (84-92) for several events on the bases and some other events. None of that is going to have much of an effect on the OBP vs. SLG analysis.

Markov Model Analysis

2001 Oakland Team Totals -- Markov Process Model					
OBP +/-	BB +/-%	Runs +/-	SLG +/-	EB +/- %	SLG/OBP
+0.010	15.00%	37	0.0213	12.20%	2.13
+0.020	31.00%	77	0.0450	25.70%	2.25
-0.010	-15.00%	-35	-0.0205	-11.70%	2.05
-0.020	-30.00%	-69	-0.0397	-22.70%	1.99
				Average:	2.10

				AL 2001 Average:	1.94
				NL 2001 Average:	1.95

▌ EB +/-% is change in extra base hits in same
proportions as season totals; total hits unchanged

Similar table to previous ones for Runs Created analysis. In this one, the ratio for Oakland is a little above 2 and that for the leagues is a little below 2, reverse of before.

Increase in scoring a little greater with higher OBP (37 vs. 35, 77 vs. 71) and the decrease is a bit smaller (-35 same, -69 vs. -70). The two models have similar ratios, but they work quite a bit differently.

2001 Oakland, Leagues

- Runs Created and Markov say an <u>extra point of OBP is worth about two of SLG</u>
- Assumes all batters the same and equal to team or league average
- What about a real lineup?
  - Different ratios by batting order position?
  - Markov model can be applied
  - Will use Oakland 2001 late season lineup

Points are pretty much self-explanatory.

Due to complexity involved and because I have not automated the computational process, I analyzed only one case, when OBP increases by 20 points for each player in the lineup in turn. Relative OBP vs. SLG values should be similar for other cases.

Lineup shown on next page is the one used for the most part during the playoffs, and these were the "regulars" at that point

Markov Model: OAK Lineup

	2001 Season:						
Player	OBP	SLG	BB +%	Runs/162 +	SA +	EB +%	SA/OBP
Damon J	0.325	0.363	35%	11.7	0.055	51%	2.80
Tejada M	0.327	0.476	48%	11.0	0.060	29%	3.02
Giambi Ja	0.483	0.660	21%	7.7	0.044	14%	2.17
Dye J	0.373	0.547	31%	8.5	0.040	16%	2.05
Chavez E	0.342	0.540	45%	8.3	0.038	15%	1.91
Giambi Je	0.393	0.450	24%	7.8	0.043	26%	2.13
Long T	0.338	0.412	40%	8.0	0.040	31%	2.05
Hernandez R	0.319	0.408	40%	8.0	0.040	26%	2.01
Menechino F	0.373	0.374	24%	8.2	0.050	38%	2.47

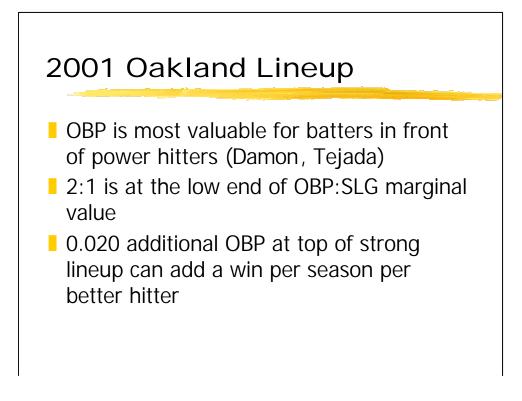Oakland 2001 late season lineup; Markov model with OBP up 0.020

OBP, SLG are full season except for Dye. His is what he did for Oakland after coming over from KC in the middle of the year. He hit better for the As, but that was consistent with his performance the prior two years.

BB+% is increase in BB for .020 increase in OBP. Some of numbers are approximations, so SA/OBP ratio may not be exactly equal to what SA on chart would indicate.

Extra OBP is more valuable in front of the power hitting portion of the lineup (exp. Jason G.), 3-5 hitters.

HR/2B/3B relations affect how much EB% is needed and how much SLG is raised.

Note low OBPs at top of order. Would be better to put Jeremy (like in 2002) and Menechino at the top.

2001 Oakland Lineup

- OBP is most valuable for batters in front of power hitters (Damon, Tejada)
- 2:1 is at the low end of OBP:SLG marginal value
- 0.020 additional OBP at top of strong lineup can add a win per season per better hitter

Usual estimate is that an extra 10 runs per season yields one more win.

0.020 increase is not unreasonable for many young players which more or less average OBP, so working on improving their strike zone judgement and discipline is worth the effort. That will probably increase their SLG also since they won't be swinging at as many pitches out of the strike zone.
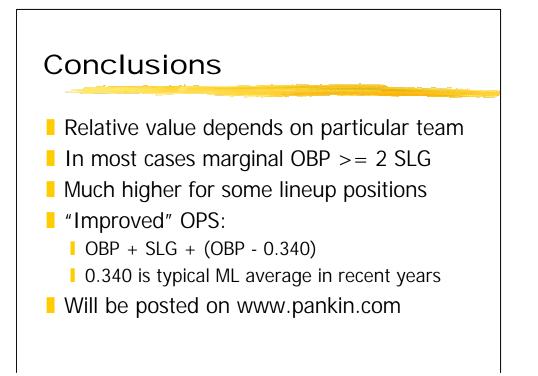
How Well Did Beane Do?

2001 Late Season Player	OBP	SLG	2002 Early Player	2001 OBP	2001 SLG	2002 OBP	2002 SLG
Damon J	0.325	0.363	Giambi Je	0.393	0.450	0.390	0.471
Tejada M	0.327	0.476	Menechino F	0.373	0.374	0.312	0.326
Giambi Ja	0.483	0.660	Hatteberg S	0.359	0.384	0.374	0.433
Dye J	0.373	0.547	Dye J	0.373	0.547	0.333	0.459
Chavez E	0.342	0.540	Chavez E	0.342	0.540	0.348	0.513
Giambi Je	0.393	0.450	Tejada M	0.327	0.476	0.354	0.508
Long T	0.338	0.412	Long T	0.338	0.412	0.298	0.390
Hernandez R	0.319	0.408	Hernandez R	0.319	0.408	0.313	0.335
Menechino F	0.373	0.374	Pena C	0.361	0.500	0.305	0.419
Average	0.364	0.470		0.354	0.455	0.336	0.428
			Difference	-0.010	-0.015	-0.027	-0.042
Markov Model Runs/G	6.1				5.8		5.1
Difference is 48.6 runs per 162 games, about 5 wins					163.6r, 16w		

Moneyball discussed thinking with Damon and Jason G. leaving and being replaced by Hatteberg and C. Pena. Loss of Jason was too much to overcome.

2001 data shown for 2002 lineup is 2001 full season except for Dye (as before) and Hatteberg, who had not played as a full season regular for Red Sox. Used his totals for 1999-2001. C. Pena had 72 plate appearances for TEX in Sept. 2001, but his performance shown was consistent with his minor league hitting.

Jeremy G. and Pena were traded during season and hit better for their new teams than for OAK. Data shown in 2002 columns is while with As. Menechino was benched for his poor hitting and had only 154 PA for the season.

Overall lineup did not hit as well as 2001 data suggested, so drop in offensive potential even higher than "planned."

Conclusions

❚ Relative value depends on particular team
❚ In most cases marginal OBP >= 2 SLG
❚ Much higher for some lineup positions
❚ "Improved" OPS:
  ❚ OBP + SLG + (OBP - 0.340)
  ❚ 0.340 is typical ML average in recent years
❚ Will be posted on www.pankin.com

We see that the marginal values depend somewhat on the particular team, and in a lineup, the values can vary quite a bit due to the strengths of the preceding and following hitters. None of this is a surprise.

Based on overall average performance, an estimate that an extra point of OBP is worth two or more of SLG seems justified. Can use 2 to improve OPS by giving a "bonus" point for each point OBP is above average or assessing a "penalty" point for each point below average.

Mention that I will post this on my web site and may write an article for By The Numbers (or BRJ?) on this topic.