By the Numbers

Volume 12, Number 4 The Newsletter of the SABR Statistical Analysis Committee November, 2002

News

"Best of BTN"

Phil Birnbaum

As you probably know, SABR has decided that 2002's third member publication will be an anthology of studies from this newsletter, entitled (no surprise) "The Best of 'By the Numbers'".

The book is almost ready – it's at the printers as I write – and it should be mailed out within a week or two of this issue going to press.

Back in July, when the book was proposed, I suggested that the book contain the best empirical studies that have run in BTN, and our members agreed, as did SABR Publications Director Jim Charlton. As a result, the book won't have any opinion pieces, book reviews, or even Charlie Pavitt's academic summaries.

In this issue

The idea is that the book should not be about sabermetrics, but about baseball. And so each study selected for the book, though using sabermetric and statistical techniques, answers an interesting question about baseball itself. Since the book is aimed at SABR members who have no sabermetric background (and, indeed, who have chosen not to join this committee), the thought was that these readers would want to learn something about the game of baseball, not about statistics.

Indeed, many of the studies have titles that directly ask the question that they answer: "What Drives MVP Voting?" (Rob Wood) and "Does Good Hitting Beat Good Pitching?" (Tom Hanrahan), to name two. These are questions that every SABR member will be interested in. We hope those who shy away from sabermetrics because of the numbers will still be intrigued enough to learn how the evidence shines light on these questions – whether or not they appreciate the statistical content.

There was no space in the book for acknowledgements, and so I will spout off a bit here. First, and most obviously, thanks to all of you who have submitted articles to BTN over the years – it is an anthology, after all, and you guys are the ones who contributed the actual content. Apologies to those who didn't make it in (especially Charlie Pavitt, whose work on hitter streakiness was originally announced as included, but had to be cut). Thanks also to our authors for proofreading their work on very tight deadline.

Second, a huge appreciative thanks to our "celebrity" contributors, professional writers and baseball men who generously contributed an introduction or two to the essays -- Paul Depodesta, F.X. Flinn, Bill James, Rob Neyer, Mat Olkin, Ron Shandler, and Jayson Stark. These contributions from baseball professionals do much to enhance the credibility of our efforts, and the enthusiasm and respect they have for our work is tremendously gratifying.

Of course, the book would not exist without the efforts of Jim Charlton, our publications director, who sold the idea to the SABR board and then spent countless hours in its editing; Glenn LeDoux, who is responsible for the book's professional look both inside and out; and John Matthew IV, who helps me edit BTN and likewise helped with the Best Of.

Thanks again, all, and we hope you enjoy the book.

Phil Birnbaum, 18 Deerfield Dr. #608, Nepean, ON, Canada, K2G 4L1, birnbaum@sympatico.ca ♦

Baseball At Altitude – SABR 33

Rod Nelson

The 33rd Annual SABR National Convention will be held July 10-13, 2003 at the Mariott City Center Hotel in Denver, Colorado. This provides an excellent opportunity for baseball's best and brightest statistical analysts to address the phenomenon that is "Baseball at Altitude". After ten years of major league play, the subject is still something of an enigma and presents a most complex challenge to the field of sabermetrics. Whether attempting to simply correlate data from games played at Mile High Stadium and Coors Field or charged with developing an organizational philosophy whose objective is to build a consistent winner, adding the altitude factor to the equation presents a formidable task like none other. The Rocky Mountain Chapter invites all SABR members to present their observations of the extraordinary game which is played in such a unique environment and the problems dealing with the disparity of home and road conditions.

The SABR33 Organizing Committee is working on several concepts for panel discussions that will be of particular interest to members of the Statistical Analysis Committee. The "Baseball at Altitude" panel will feature a cross-section of baseball experts on physics, statistics, history, and player personnel. Others intriguing topics include "The Relief Pitcher and the Hall of Fame", "Baseball Simulation Games" and others.

Members of the Statistical Analysis Committee are invited to develop presentations or poster displays of their work with the possibility for post-convention publication of papers on the "Baseball at Altitude" theme. The Presentations Chair is Gail Rowe (growes36@attbi.com). Watch SABR-L and the SABR Bulletin for additional details.

Submissions

Phil Birnbaum, Editor

Submissions to *By the Numbers* are, of course, encouraged. Articles should be concise (though not necessarily short), and pertain to statistical analysis of baseball. Letters to the Editor, original research, opinions, summaries of existing research, criticism, and reviews of other work (but no death threats, please) are all welcome.

Articles should be submitted in electronic form, either by e-mail or on PC-readable floppy disk. I can read most word processor formats. If you send charts, please send them in word processor form rather than in spreadsheet. Unless you specify otherwise, I may send your work to others for comment (i.e., informal peer review).

If your submission discusses a previous BTN article, the author of that article may be asked to reply briefly in the same issue in which your letter or article appears.

I usually edit for spelling and grammar. (But if you want to make my life a bit easier: please, use two spaces after the period in a sentence. Everything else is pretty easy to fix.)

If you can (and I understand it isn't always possible), try to format your article roughly the same way BTN does, and please include your byline at the end with your address (see the end of any article this issue).

Deadlines: January 24, April 24, July 24, and October 24, for issues of February, May, August, and November, respectively.

I will acknowledge all articles within three days of receipt, and will try, within a reasonable time, to let you know if your submission is accepted.

Send submissions to:

Phil Birnbaum
18 Deerfield Dr. #608, Nepean, Ontario, Canada, K2G 4L1
birnbaum@sympatico.ca

Academic Research: Motivations for Integrating, and Yet Another Ranking Stat

Charlie Pavitt

The author describes two current academic studies. In the first, researchers investigate whether the best or worst teams were quickest to integrate. In the second, it's another method for ranking players – "heavy on the math but light on the thinking."

This is one of a series of reviews of sabermetric articles published in academic journals. It is part of a project of mine to collect and catalog sabermetric research, and I would appreciate learning of and receiving copies of any studies of which I am unaware. Please visit the 2002 edition of the Statistical Baseball Research Bibliography at its new location www.udel.edu/communication/pavitt/biblioexplan.htm . Use it for your research, and let me know what is missing.

Brian L. Goff, Robert E. McCormick, and Robert D. Tollison, Racial Integration as an Innovation: Empirical Evidence from Sports Leagues, American Economic Review, Volume 92 Number 1, March 2002, pages 16-26

This one provides another new angle on the racial discrimination issue, asking whether the first teams to integrate were the weakest (presuming they would have the most to gain from recruiting black players) or the strongest (presuming they were the best managed and thus more quickly realized the advantages black players would provide). Using data from 1947 through 1971, the authors found that the more successful teams in one season (as measured by Games Behind in the final league standings) had a relatively greater number of black players during the next season than the less successful teams. This tendency was stronger in the National League, with the Dodgers and Giants showing the way and the other NL teams apparently following their lead in order to compete. The effect in American League was weaker; recall that the Yankees were late integrators. The rate of integration followed the normal S-shaped function for adoptions of innovations of any type; slowly at first as only the more adventurous are willing to try something new, then quicker as others join in, and finally slower again as the laggards slowly capitulate. Goff et al. found a similar effect in ACC basketball using the 1965-66 through 1991-92 seasons.

Gary Koop, Comparing the Performance of Baseball Players: A Multiple-Output Approach, Journal of the American Statistical Association, Volume 97 Number 459, September 2002, pages 710-720

What we have here is another attempt at developing an overall method for ranking players, heavy on the math but light on the thinking. If I understand it correctly, which I admit I might not, Koop uses various ranking methods, but highlights one in which performance is estimated by equally weighing three indices; singles plus walks plus hit by pitch per plate appearance, doubles plus triples per plate appearance, and home runs per plate appearance. Given how few triples there are, the second index is really doubles per plate appearance, evaluated as equally important to home runs. Therefore it is not surprising to find, for the years 1995 through 1999, Edgar Martinez ranked first, Mark Grace ranked fourth, and Chuck Knoblauch ranked tenth. Perhaps Koop's method has potential (I leave that judgment to the better methodologists among us), but the present effort is substantively useless.

Charlie Pavitt, 812 Carter Road, Rockville, MD, 20852, chazzq@udel.edu ♦

Does Experience Help in the Postseason?

Tom Hanrahan

Do players perform better in the World Series if they already have postseason experience? Many baseball people assume the answer is yes – but here, the author examines the evidence and concludes otherwise.

Do baseball players fare better in the postseason when they have some relevant postseason experience behind them? My research says the answer is a clear *no*. Managers' efforts to build teams with players who "have been there before" appear to be fruitless ventures, sacrificing money and possibly quality for no apparent gain.

How to measure postseason experience?

Common baseball wisdom says that one factor in achieving postseason success is having postseason experience. The thinking is that because the World Series is such a unique event, that those players who have not been there before are more likely to fall victim to nervousness, lack of confidence, or other conditions that would adversely affect their play. This sounds like a reasonable theory. The success of the Yankees in recent postseason play would be one piece of positive evidence, but then this last World Series had young Lackeys and Rodriguez-es all over it.

The big question is, how do we tell if the theory is true? We could analyze by team or individual success. From a team standpoint, we could compare the winning percentage of teams who have played in the playoffs more recently than their opponents. A difficulty with this method would be controlling for the quality of the team, particularly the different mix of players from year to year. So, I chose to focus on individual performance. The obvious decision here is looking at hitters or pitchers. Here I determined that using pitchers would be a far better choice, for many reasons:

1. Ease of choosing a measuring stick. For batters, we can use some standard measure such as OPS, but in short series, clutch hitting can be more important than overall performance, so one could easily argue that R and RBI, broken down into game situations, are the best measure. For pitchers, ERA seems to be the obvious best measure, and it is easy to obtain.

2. Sample size. In a playoff series, starting pitchers often face 70 batters or more. Hitters rarely get more than 30 plate appearances in series 7 or fewer games.

3. Pressure. I decided to use pitchers who started games, because a pitcher named by his manager to take the mound is obviously the focal point on the field for his team that day (exception: Barry Bonds in 2002).

What data to use?

The postseason has changed over the years; from the World Series only, to two and now three sets of playoffs. Because of the difficulty in defining "postseason pressure" or "experience" with mixed sets, I decided to only use pitchers who started World Series games, and whose first WS start was prior to 1969, the year divisions were created. This yielded 65+ years of WS play.

I created two sets of data for each pitcher. The data from "inexperienced" hurlers were those innings thrown by pitchers who made at least one start in their first WS appearance. In other words, if a pitcher first appeared in a WS as a reliever only, he was ineligible; this was done so I did not have to decide whether a relief appearance counted as "experience". Then I used his combined total of all other WS innings as "experienced" data, as long as he made at least one more start in a succeeding WS. Pitchers who did not start any games in any successive WS were again not used.

I was surprised when collecting the data that there are many discrepancies among sources for pitchers' ERAs in the early part of the century. I eventually chose the baseball1.com database as my standard reference.

And the results are……

This method yielded 97 pitchers, who combined for 3710 innings pitched in WS play. Fifty-four pitchers (56%) had a lower ERA in their initial Series appearance than in future games. Forty-two pitched better when they were experienced in Series play. One had identical ERAs; Dutch Leonard threw 9 IPs of 1 run ball in both cases. Raw data totals:

First WS IP	First WS ERA	Future WS IP	Future WS ERA	ERA difference
1213	2.69	2497	2.97	+0.28

Possibly a more accurate way of determining the "experience effect" would be to create matched sets of data, rather than combining all of the data into one pool. As an example, suppose the entire data set consisted of only 2 pitchers. Their mythical totals are listed below:

	First WS		Future WS		
Hurler	IP	ERA	IP	ERA	ERA difference
A	5	2.00	25	3.60	+1.60
B	25	4.00	5	4.00	0.00
Combined	30	3.67	30	3.67	0.00 (!)

Pitcher A was much worse in future games, B was the same; yet overall, there is no difference! This is because of the disparity in innings and in performance; A was better overall, or possibly pitched in an era when it more difficult to score runs. Clearly, it is possible that in the instance above, one could interpret the combined results to say that the pitchers were better in their initial WS appearances.

To combat this situation, I created matched sets of data for each pitcher, weighting the difference in ERA by how many innings each pitcher threw. I used the harmonic mean of each pitcher's first and future innings as the weight for each pair. Here is the data for the first 3 pitchers alphabetically:

	First WS		Future WS		ERA	weighted	ERA diff
	IP	ERA	IP	ERA	difference	IP	times IP
V Aldridge	18	4.42	8	7.15	+2.73	11.1	30
G Alexander	18	1.53	25	5.02	+3.49	20.9	73
E Auker	11	5.56	6	3.04	-2.52	7.8	-20
Combined	47	3.58	39	5.15	+2.11*	39.8	83.7

Aldridge's weighted IP are found by 2 / (1/18 + 1/8) = 11.1. The ERA difference and weighted IP are multiplied for each pitcher to get the right-most column. 83.7 divided by 39.8 yields a composite ERA difference of 2.11. The pitcher with the most weighted innings in WS play is Christy Mathewson, who famously tossed 27 shutout innings in his initial 1905 appearance, and then threw 75 more innings with an ERA of 1.44 later in his career.

When combined using this method, there is a total of 1374 weighted IP. Pitchers as a whole had an ERA that is a full 0.64 runs per game *higher* in their "experienced" Series play than in their first outings.

While not being beyond the bounds of statistical significance, these results clearly show no apparent advantage of postseason experience in the first two-thirds of the 20th century.

Further investigation could be performed to attempt to account for age differences, park effects, or other factors, but I reason that the chance of these other influences is very unlikely to be large enough to sway the results of this study. Experienced pitchers have done no better in the pressure cooker of the Series than first-timers. Tell that to your favorite announcer.

Notes: Rob Wood was kind enough to review my data, correct lots of missing or bad info, and pointed out the data differences in various sources. Raw data for the 97 pitchers in the study are available on request.

Tom Hanrahan, 21700 Galatea St., Lexington Park, MD, 20653, HanrahanTJ@navair.navy.mil. ♦

Get Your Own Copy

If you're not a member of the Statistical Analysis Committee, you're probably reading a friend's copy of this issue of BTN, or perhaps you paid for a copy through the SABR office.

If that's the case, you might want to consider joining the Committee, which will get you an automatic subscription to BTN. There are no extra charges (besides the regular SABR membership fee) or obligations – just an interest in the statistical analysis of baseball.

To join, or for more information, send an e-mail (preferably with your snail mail address for our records) to Neal Traven, at beisbol@alumni.pitt.edu. Or write to him at 4317 Dayton Ave. N. #201, Seattle, WA, 98103-7154.

Receive BTN by E-mail

You can help save SABR some money, and me some time, by receiving your copy of *By the Numbers* by e-mail. BTN is sent in Microsoft Word 97 format; if you don't have Word 97, a free viewer is available at the Microsoft web site (http://support.microsoft.com/support/kb/articles/Q165/9/08.ASP).

To get on the electronic subscription list, send me (Phil Birnbaum) an e-mail at birnbaum@sympatico.ca. If you're not sure if you can read Word 97 format, just let me know and I'll send you this issue so you can try

If you don't have e-mail, don't worry – you will always be entitled to receive BTN by mail, as usual.

Should Teams Walk or Pitch to Barry Bonds?

Jerome P. Reiter

Is Barry Bonds so productive that it's often better to walk him than to let him hit away? Here, the author investigates when, if ever, the IBB is called for against Bonds, by analyzing play-by-play data from 2001-2002.

In 2001, Barry Bonds of the San Francisco Giants had arguably the greatest individual season in the history of major league baseball. He set the record for home runs in a season with 73. He hit for the highest slugging percentage ever at .863, breaking Babe Ruth's 1920 mark of .847. He knocked in 137 runs, good for fourth best in the National League, and his batting average was .328, good for seventh best in the National League. Bonds achieved these gaudy statistics despite being walked by opposing pitchers a major league record 177 times, besting Ruth's 1923 record of 170. Not surprisingly, Bonds was voted the Most Valuable Player in the National League by an overwhelming margin. It was the fourth time he won the MVP award.

Bonds did not slow down in 2002. He began the campaign with an incredible display of power, hitting five home runs in the Giants' first four games. His year-end statistics were remarkable: a major league leading .370 batting average, a National League second-best 46 home runs, and a National League sixth-best 110 runs batted in. He reached base in 58.2% of his plate appearances, eclipsing the previous record of 55.1% set by Ted Williams in 1941. He passed Frank Robinson for fourth all time in career home runs, with only Willie Mays, Babe Ruth, and Hank Aaron ahead of him. Perhaps most amazingly, Bonds finished with 198 walks, shattering the record he set just last season. Once again, Bonds was voted MVP in the National League.

The inflated walk totals in 2001 and 2002 reflect increased use of a strategy for dealing with Bonds' awesome power, namely not to let him have a chance to hit the baseball. Why risk pitching to a player who averaged one home run every 6.5 at bats in 2001 and every 8.7 at bats in 2002? However, walking Bonds is not failsafe. Putting him on base via walk could actually help the Giants, since having runners on base greatly increases a team's chance of scoring runs. Plus, as prolific a batter as he is, over his career Bonds has made an out roughly 70% of the times he is not walked. Why not pitch to him since outs are the most likely outcome?

Thus, we are confronted with an interesting question of baseball strategy: is it better to walk Barry Bonds or to pitch to him? Clearly, the answer depends on the game situation when Bonds steps to the plate. For example, walking Bonds has different consequences when there are zero outs and runners on first and second as opposed to when there are two outs and a runner on third. The answer also depends on the outcomes the opposing manager is concerned about. The manager who seeks to prevent even a single run—for example in late or extra inning situations in which one run can result in a lost game—evaluates walks differently than does the manager who will concede one run to reduce the possibility of multiple runs.

In this article, we examine data from the 2001 and 2002 seasons to investigate if and when it is better strategy to walk rather than pitch to Barry Bonds. We focus primarily on two game situations: when there is no one on base, and when there is a player on first base only. The conclusions suggested by the data are somewhat surprising: in these two situations, walking Bonds is not more effective at preventing runs than letting him hit. In fact, the data even suggest—very weakly—that it is preferable in some situations to let Bonds swing away.

Data Used in Analyses

To assess the two strategies, we examine data on Bonds' plate appearances in the 2001 and 2002 seasons. The data were retrieved from the web site http://www.cbs.sportsline.com, which has links to pitch-by-pitch game logs for each game in 2001 and 2002. Data for a few games were unavailable because of invalid web links; these games are excluded from the analyses. This should not skew results since these games are missing completely at random, that is, they are missing for reasons unrelated to the variables measured.

Some Characteristics of Bonds' Plate Appearances

There are twenty-four possible game situations when a player steps into the batter's box. These are obtained by crossing the three possible out values and the eight possible configurations of players on base. Table 1 displays the number of walks and plate appearances by Bonds in 2001 and 2002 in each of these twenty-four game situations. In this table, and for all analyses, intentional and unintentional walks and hit by pitches are included in the walk totals.

In most games in 2001, Bonds batted third in the San Francisco Giants batting order, so that he often came up with two outs and the bases empty or with one out and a runner on first base. In 2002, Bonds typically batted fourth in the order, often leading off innings or batting with runners on base and two outs. In both seasons, most of Bonds' walks were issued with two outs. This is understandable: with two outs, some managers feared Bonds' ability to get extra-base hits more than his teammates' ability to score him from first base. Managers were more reluctant to walk Bonds with zero or one outs, perhaps because multiple players have the chance to advance him home. Bonds walked most frequently when runners were in scoring positions and first base was unoccupied. Walking players in these situations is a common baseball strategy.

Game situations other than those in the "None On" and "First Only" categories have few observations. Hence, in the analyses that follow, we focus primarily on the None On and First Only categories.

Bonds is in the heart of the Giants batting order, so that we expect innings in which he appears at the plate to be the most productive for the Giants. In our analyses, we are not concerned with the total number of runs in each inning, but rather the total number of runs scored in the inning after the first pitch to Bonds. Hence, for purposes of analyses, we redefine runs in an inning to be the number of players crossing the plate in the inning after the first pitch to Bonds in that inning. Figure 1 shows the frequencies of the run totals for all innings in which Bonds stepped to the plate. In more than 60% of the innings, the team scores zero runs. Scoring more than two runs in an inning is rare. This suggests that fear of large run totals should not be a strong factor in the decision to walk or pitch to Bonds.

		None On	First Only	Second Only	Third Only	First + Second	First + Third	Second + Third	Bases Loaded
0 out	2001	20/111	3/30	1/9	0/1	1/14	1/7	1/3	0/1
	2002	32/116	7/25	2/4	1/2	2/9	1/3	2/2	0/3
1 out	2001	17/88	19/84	12/23	3/12	5/15	2/5	2/2	0/3
	2002	22/86	13/45	15/21	12/13	4/18	2/7	4/4	0/4
2 out	2001	41/151	14/34	13/18	5/7	9/15	4/9	1/1	1/5
	2002	36/117	16/48	14/21	1/6	5/15	7/11	1/1	0/4

Table 1: Bonds' walks / plate appearances in 2001 and 2002

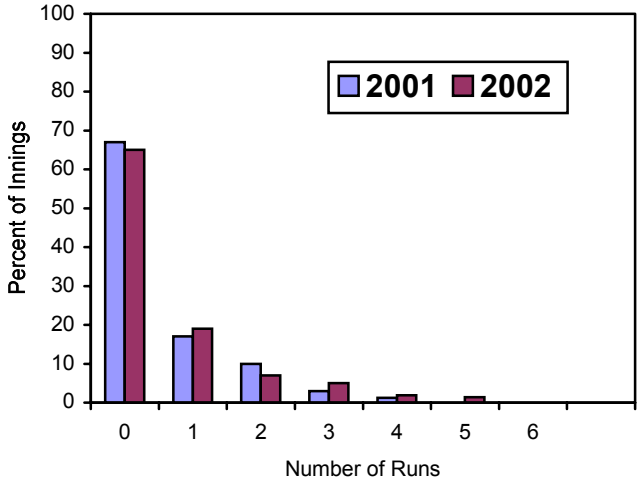

Figure 1: Percentages of number of runs in innings when Bonds appears at the plate

Walking Bonds To Prevent Giants From Scoring At Least One Run

When games are close in late or extra innings, the goal is not to let in any runs. This goal motivates the first question we address: does walking rather than pitching to Barry Bonds reduce the chance that at least one run scores?

For some game situations, baseball strategy dictates that walking Bonds is the smart choice. These include the situations in which first base is unoccupied and there are runners in scoring position (on second or third base). Consider the situations on an out-by-out basis. With two outs, walking Bonds is preferable since Bonds has a higher batting average than those players batting immediately after him. That is, Bonds is more likely to get a hit, thereby driving home the run, than those other players. With one out, walking Bonds sets up a double play that potentially can end the inning without any runs scoring. With zero outs, walking Bonds sets up force plays that prevent the lead runner from advancing further.

For other situations, it is clear that a walk is not the smart choice. When the bases are loaded, a walk automatically gives the Giants a run. With runners on first and second and less than two outs, a walk advances a runner to third base where he can score on a fly out or a well-placed ground out.

For the remaining situations, it is not clear from baseball strategy whether walking Bonds is the smart choice. Hence, it is useful to examine data for evidence of the success of one strategy over the other. Figure 2 displays the percentages of innings in which the Giants score at least one run after Bonds comes to the plate for the None On and First Only situations. Within each situation, the left bar shows the percentage when Bonds is walked, and the right bar shows the percentage when Bonds is pitched to. Each bar represents the combined percentage obtained by pooling both years of data. The annual percentages are above the bars, with the 2001 percentage shown on top of the 2002 percentage. For example, in 2001 the Giants scored in 6 of the 20 innings in which Bonds was walked with none on and no outs, and in 2002 the Giants scored in 17 of 32 such innings. Thus, the Giants scored when Bonds was walked with none on and no outs a combined (6+17)/(20+32) = 44% of the time.

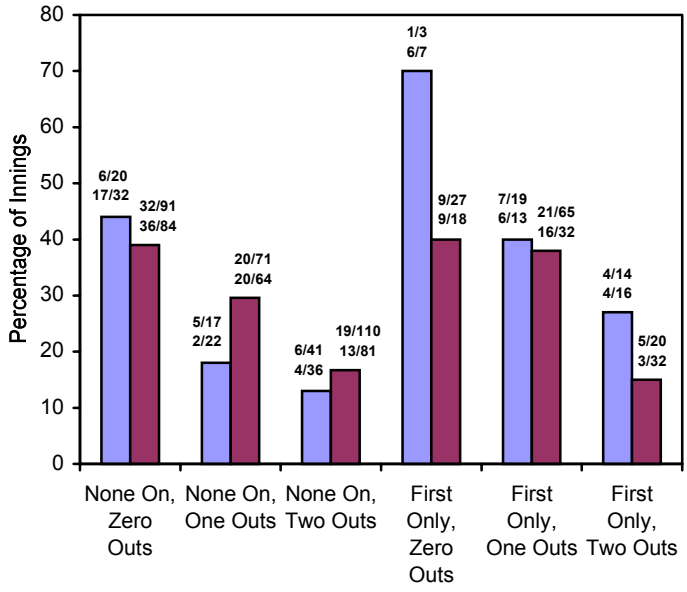

Figure 2: Percentage of innings Giants score at least one run after first pitch to Bonds. The left bars are the innings in which Bonds walked; the right bars are when he was pitched to. The fractions above each bar are the scoring innings / total innings, for 2001 and 2002 respectively.

Pooling the data from 2001 and 2002 simplifies comparisons of the strategies. Additionally, the combined percentages are based on larger numbers of innings than the annual percentages, which improves our ability to differentiate the effectiveness of the strategies. A drawback to pooling the data is that it masks any differences across years.

At first glance, the combined percentages suggest competitive advantages for each strategy. With none on and at least one out, walking Bonds seems more effective than pitching to him. This may be because in these situations, the risk that Bonds hits a home run outweighs the risk that he scores when put on first base. With none on and no outs, walking Bonds seems less effective than pitching to him. This suggests that avoiding Bonds' home run power is outweighed by beginning an inning with a free pass. With one man on base, pitching to Bonds seems to be the better strategy. The walk advances the runner on first to scoring position, and it may be that the risk of that runner scoring outweighs the risk of Bonds driving in the runner from first.

These percentages are based on a limited number of plate appearances. Suppose there is no difference in the true probabilities of the Giants scoring when Bonds walks or hits. Could these apparent differences be plausibly explained by random chance? To answer this question, we conceive of a hypothetical population of Bonds' plate appearances under the same conditions that existed in 2001 and 2002, and we consider the plate appearances in 2001 and 2002 a random sample from this hypothetical population. Under this framework, the answer to our question is yes. When we combine the 2001 and 2002 data, the p-values for all six two-tailed statistical hypothesis tests are greater than 0.10. In other words, if walking and pitching to Bonds are equally effective, in each situation we expect to see differences in the combined percentages as large (or larger) than those in 2001 and 2002 more than 10% of the time.

More plate appearances would help us eliminate random chance as a plausible explanation for the differences in the run percentages, and thereby help us determine if the apparent advantages of the strategies evident in the 2001 and 2002 season are real. However, the 2001 and 2002 data, while suggestive, do not distinguish clear favorites. In these situations, walking and pitching to Bonds may well be similar in effectiveness.

Before accepting this conclusion, we must be sure that our comparisons are fair to both strategies. That is, up to the point when Bonds steps to the plate, the innings in which Bonds walks should have similar characteristics to those in which he hits. In general, when comparing two strategies, fairness can be assured by assigning the strategies to the experimental units at random. We do not have this set-up when analyzing the 2001 and 2002 data. Decisions to walk or pitch to Bonds were made by managers rather than at random. Hence, we have to check whether the walk-innings differ from the hit-innings in ways that could affect the chances of scoring runs.

Other than game situation, the primary variable that affects the number of runs scored is the opposing pitcher's quality. Weak pitchers are likely to give up more runs than strong pitchers, regardless of whether they walk or pitch to Bonds. A pitcher's quality can be measured by his earned run average (ERA) over his career, which roughly equals the total number of all runs allowed by the pitcher divided by the number of innings he has pitched. Based on investigations within each game situation, the distributions of ERA are similar for the innings Bonds walks and innings he does not walk in both 2001 and 2002. Hence, any effects of ERA on runs are equally present in the side-by-side percentages of Figure 2. The comparison is fair with respect to ERA.

Other background characteristics we measure include the player hitting after Bonds, whether the game is in San Francisco or at other stadiums, the score, the inning, and the game number in the season. In roughly 85% of his at bats in 2001, Bonds was followed by Jeff Kent, and walks were similarly distributed in games when Kent or someone else batted after Bonds. In 2002, Kent followed in roughly 52% of plate appearances; Benito Santiago followed in roughly 28% of plate appearances; Reggie Sanders followed in roughly 18% of plate appearances; and, other players accounted for the remaining 2%. In 2002, Bonds walked in roughly 33% of his at bats when followed by Kent, 33% when followed by Santiago, and 37% when followed by Sanders. These percentages are similar enough that the comparisons of walk-innings and hit-innings should not be affected greatly by differences in the players batting after Bonds. For the other variables, there are one or two game situations for which the variables' distributions differ in the walk-innings and the hit-innings. These differences should have minimal effect on the comparisons, because these variables do not have strong relationships with the probability of scoring runs. Players of Bonds' caliber, and those who bat after him, try equally hard to score runs regardless of the stage of the game and the time of the season.

Hence, our conclusions stand. The data suggest walking Bonds may be preferable when there is no one on and at least one outs, and pitching to him may be preferable when there is a runner on first base or when Bonds leads off an inning. But, these are only suggestions; we cannot rule out chance variation as a plausible explanation of these suggestions. We ultimately conclude that, when the objective is to prevent any runs from scoring, the data do not provide sufficient evidence of a substantial difference in the effectiveness of walking versus pitching to Barry Bonds.

Does Walking Bonds Help Avoid Big Innings?

Walking Bonds may not have much effect on the chance of scoring at least one run, but does it have an effect on the chance of scoring runs in general? That is, do we expect bigger run totals when Bonds is walked as opposed to when he hits? Let's use the data to investigate this question for the None On and First Only game situations.

Let the variable y represent the number of runs scored in an inning. Ideally, we'd know the probability the Giants score y runs for each value of y in each game situation under both strategies, so that we could compare the two strategies in any situation by comparing their probabilities. Of course, we don't know these probabilities, so we use the 2001 and 2002 data to learn about them.

It is relatively rare that $y>3$. In fact, $y>3$ for only 1.2% of all innings in the None On situation and 2.9% of all innings in the First Only situation. This suggests that we can simplify analyses without losing much information by collapsing runs into four categories: $y=0$, $y=1$, $y=2$, and $y\geq 3$.

Natural estimates of the probabilities for these four run categories are the proportions of runs that fall in each category in Bonds' combined plate appearances from 2001 and 2002. These proportions are displayed in Table 2. The table also includes the averages and standard deviations of the number of runs. For positive run categories, there is an interesting trend in the data. Let x be the number of runners on base when Bonds steps to the plate. When Bonds hits, the Giants appear more likely to score exactly $(x+1)$ runs than when he walks. On the other hand, when Bonds walks, the Giants appear more likely to score $(x+2)$ or more runs than when he hits. The averages of runs are typically as small or even smaller when Bonds hits than when he walks, the one exception being when the bases are empty with one out. This exception is explained by the unusually low number of runs scored in 2002 in this situation, only two runs in twenty-two innings. Overall, these patterns suggest that it might be preferable to pitch to Bonds rather than to walk him in the None On and First Only situations, although opposing teams face risks when using either strategy.

Situation	Year	No Walk						Walk					
		N	0	1	2	3+	Avg. (SD)	N	0	1	2	3+	Avg. (SD)
None on, Zero outs													
	2001	91	.65	.25	.09	.01	.49 (.9)	20	.70	.20	.10	.00	.40 (.7)
	2002	84	.57	.25	.11	.07	.71 (1.1)	32	.47	.22	.22	.09	1.03 (1.3)
	Comb.	175	.62	.25	.10	.03	.60 (1.0)	52	.56	.21	.17	.06	.79 (1.1)
None on, One outs													
	2001	71	.72	.21	.06	.01	.39 (.9)	17	.70	.12	.12	.06	.53 (.7)
	2002	64	.69	.23	.03	.05	.48 (.9)	22	.91	.05	.00	.04	.18 (.7)
	Comb.	135	.71	.22	.05	.02	.43 (.9)	39	.82	.08	.05	.05	.33 (.8)
None on, Two outs													
	2001	110	.83	.17	.00	.00	.17 (.5)	41	.85	.05	.08	.02	.26 (.9)
	2002	81	.84	.12	.04	.00	.20 (.5)	36	.89	.05	.03	.03	.25 (.9)
	Comb.	191	.83	.16	.01	.00	.18 (.5)	77	.87	.05	.06	.02	.26 (.9)
First only, Zero outs													
	2001	27	.67	.07	.15	.11	.74 (1.2)	3	.67	.00	.00	.33	2.00 (3.5)
	2002	18	.50	.33	.06	.11	.83 (1.1)	7	.14	.43	.14	.29	1.86 (1.7)
	Comb.	45	.60	.17	.11	.11	.78 (1.2)	10	.30	.30	.10	.30	1.90 (2.1)
First only, One outs													
	2001	65	.68	.14	.15	.03	.72 (.9)	19	.63	.16	.05	.16	.79 (1.3)
	2002	32	.50	.25	.12	.13	1.00 (1.4)	13	.54	.23	.08	.15	.85 (1.1)
	Comb.	97	.62	.18	.14	.06	.81 (1.1)	32	.59	.19	.06	.16	.81 (1.2)
First only, Two outs													
	2001	20	.75	.00	.20	.05	.55 (1.0)	14	.71	.07	.07	.15	.78 (1.5)
	2002	32	.91	.06	.03	.00	.13 (.4)	16	.75	.13	.06	.06	.56 (1.2)
	Comb.	52	.85	.04	.09	.02	.29 (.7)	30	.73	.11	.06	.10	.67 (1.3)

Table 2: Total innings (N), proportions for run categories, and averages and standard deviations of runs scored (2001, 2002, and both years combined)

As before, we should consider random variation when interpreting these sample proportions and sample averages. Suppose there is no difference in the effectiveness of walking or pitching to Barry Bonds in reality. Could the differences between walk-innings and hit-innings observed in 2001-2002 be plausibly explained by random chance? Let's again conceive of Bonds' combined 2001-2002 plate appearances as a random sample from a hypothetical population of his plate appearances under the current conditions in the league. We seek to learn about the differences in average runs in this hypothetical population when walking versus pitching to Bonds. For the game situations with a runner on first only and zero outs, and a runner on first only and two outs, the p-values for two-tailed statistical hypothesis tests are very close to .10, a small enough value to cast doubt on chance error as an explanation for the differences in the sample averages in these situations. For these situations, the data provide weak evidence that favors pitching to Bonds. For the other game situations, the p-values of the two-tailed statistical hypothesis tests are all substantially greater than .10, making it hard to rule out chance errors as explanations of the differences in the sample averages. For these other situations, there is not enough evidence to determine conclusively that one strategy results in fewer runs on average than the other strategy does.

Concluding Remarks

There has been an incredible number of walks issued to Barry Bonds in the last two years. Given his prodigious home run power, it is understandable why managers fear pitching to him. However, the data from 2001 and 2002 suggest that there is little difference in opposing teams' ability to prevent runs when walking Bonds versus when letting him hit. In fact, the data suggest that it may be better to pitch to Bonds than to walk him in some game situations. Personally, I'm rooting for managers to pitch to Bonds in 2003, so that I can enjoy another record-breaking season!

Jerome P. Reiter, Institute of Statistics and Decision Sciences, Box 90251, Duke University, Durham, NC 27708-0251,
jerry@stat.duke.edu ♦

E-mail Changes

If you normally receive "By the Numbers" by e-mail, but you found this issue in your physical mailbox instead, it's probably because your e-mail address changed. If you'd like to switch back to an e-mail BTN, please drop me (Phil) a line with your new e-mail address, and I'll switch you back to the electronic version. I'm at birnbaum@sympatico.ca.

Informal Peer Review

The following committee members have volunteered to be contacted by other members for informal peer review of articles.

Please contact any of our volunteers on an as-needed basis - that is, if you want someone to look over your manuscript in advance, these people are willing. Of course, I'll be doing a bit of that too, but, as much as I'd like to, I don't have time to contact every contributor with detailed comments on their work. (I will get back to you on more serious issues, like if I don't understand part of your method or results.)

If you'd like to be added to the list, send your name, e-mail address, and areas of expertise (don't worry if you don't have any - I certainly don't), and you'll see your name in print next issue.

Expertise in "Statistics" below means "real" statistics, as opposed to baseball statistics - confidence intervals, testing, sampling, and so on.

Member	E-mail	Expertise
Jim Box	im.box@duke.edu	Statistics
Keith Carlson	kcarlson2@mindspring.com	General
Rob Fabrizzio	rfabrizzio@bigfoot.com	Statistics
Larry Grasso	l.grasso@juno.com	Statistics
Tom Hanrahan	HanrahanTJ@navair.navy.mil	Statistics
John Heer	jheer@walterhav.com	Proofreading
Keith Karcher	kckarcher@compuserve.com	General
Chris Leach	chrisleach@yahoo.com	General
John Matthew IV	john.matthew@rogers.com	Apostrophes
Duke Rankin	RankinD@montevallo.edu	Statistics
John Stryker	johns@mcfeely.interaccess.com	General
Dick Unruh	runruhjr@dtgnet.com	Proofreading
Steve Wang	scwang@fas.harvard.edu	Statistics