BASEBALL ANALYST

JOURNAL OF SABERMETRICS
April, 1988 Vol. #35

I promised to get the April issue out in July, and it's July as I write this, so I'm not late by my own standards, right? Well, I'm catching up as fast as I can.

Six articles here after the intro:

I've been writing a lot of stuff here not because I am short of material or bored and have nothing to do but, I suspect, because with the end of the Abstract I have nowhere else to go with a sabermetric article; a year ago I would have filed the two short articles here and worked them into the book.

I do need material, of course--I've got to keep popping out issues at one a month for awhile. If you've got anything, I need to see it. There is a suggestion for an article on Page 15.

Editor and Publisher: Bill James Address: Box 171, Winchester, KS 66097

The Effect of Being Traded on Batting Performance: More Academic Baseball Research

Charlie Pavitt

Baseball research gets published in the darndest places. Back in Analyst 27, I reported on three studies in the social science literature on the effect of managerial change on team performance. Well folks I'm back again, this time with two studies on the effect of being traded on batting performance. As is usually the case with social science research, the studies were actually concerned with a larger issue; in this case, the general question of how job changes affect worker performance. Baseball was used as an example because performance criteria for baseball is far clearer than for most professions (we know how a averages) and data is readily available.

In order to understand the theoretical rationale for these studies, I need to discuss what is known about the effect of arousal on task performance. When performing a task, a person will be undergoing a certain amount of arousal (nervousness, stress, excitement) at any given time. High levels of arousal affect performance, but how it affects performance depends on how well the person has learned the task. If a person is excited or nervous and unfamiliar with the task, their performance tends to be worse than normal; but if an aroused person is familiar with the task, their performance tends to be better than normal. Public speaking is a good example of this; the well-rehearsed speaker can channel their nervousness into good performance while the ill-rehearsed speaker panics.

Bateman, Karwan, and Kazee (1983) apparently assumed that batting is a well-learned skill among major leaguers, because they hypothesized that greater arousal would lead to better performance. Most simply, they believed that being traded leads to greater arousal among players. If traded between seasons, the increase in arousal has time to dissipate before the next season, so performance should not improve. If traded during a season, the increase in arousal should lead to better performance for the rest of the season. Further, as players are more familiar with their own league, the task of hitting is more well-learned for players traded within leagues than for players traded across leagues. As a consequence, performance increases should be greater following trades within a league.

Bateman et al. analyzed BA, HR per AB, and RBI per AB, for 97 players who had been traded or sold (no free agent movers) between 1976 and 1980 and had at least 50 AB both before and after the trade. They found clear support for their first hypothesis. Within-season trades led to increases in BA (.240 to .269), RBI/AB (.108 to .133) and HR/AB (.019 to .030), while between-season trades did not (BA, .259 to .251; RBI/AB, .111 to .114; HR/AB, .019 to .018). They only found partial support for their second hypothesis. Same-league trades led to improved performance (BA; .243 to .262; RBI/AB, .111 to .135; HR/AB, .023 to .026), but so did across-league trades, although the increases were smaller (BA; .254 to .258; RBI/AB, .109 to .117; HR/AB, .017 to .023).

Jackson, Buglione and Glenwick (1988) limited their study to

within-season trades. In contrast with Bateman et al., they assumed that batting is a poorly-mastered skill, given the fact that batters succeed on getting on base less than half of the time. It follows from this view that arousal would hurt, rather than help, performance. Jackson et al. apparently based their hypotheses on an interview with Bobby Murcer concerning player's feelings about being traded. Normal times in a player's life lead to normal arousal and normal performance. When rumors of being traded are rampant, the player gets nervous (aroused), which lowers performance, which increases the odds of being traded in a self-fulfilling prophecy. A trade decreases arousal below normal because the player feels welcome and wanted by their new team, increasing performance above normal for the rest of the season. Next season brings a return to normal arousal and performance. Complicating matters is the player's amount of experience. The more experienced the player is, the less he is affected by being traded, so the less variation there is in his performance.

Jackson et al. studied BA and SA for 59 players who were traded between 1964 and 1981 and had at least 75 AB during the three years previous to the trade, at least 75 AB during the year of the trade before the trade occurred (the "pretrade period"), at least 75 AB during the year of the trade after the trade occurred (the posttrade period"), and at least 75 AB the following year. Their findings were consistent with their first hypothesis. Batting performance decreased in the pretrade period from the previous year, supposedly as a consequence of the rumors (BA, .260 to .235; SA, .390 to .340; all figures are approximate as I had to read them off a graph rather than a table of numbers), increased during the posttrade period (BA, .270; SA, .390), then decreased the next season (BA, .255; SA, .375). Their second claim had mixed support. Experience was not related to pre or posttrade performance changes, but was related to changes in performance between the post-trade period and the following season in the case of BA only (high experience, .271 to .272; average experience, .266 to .258; low experience, .261 to .241).

What is clear from these studies is that within-season trades help performance for the rest of the season. What is unclear is why. The two articles suggest opposing reasons for this increase. For Bateman et al., being traded increases arousal, and arousal helps performance. For Jackson et al., being traded decreases arousal, and arousal hurts performance. Given the fact that they based their claims on their interview with Murcer, I'll side for now with the latter claim.

References

Bateman, T. S., Karwan, K. R., & Kazee, T. A. (1983). Getting a fresh start: A natural quasi-experimental test of the performance effects of moving to a new job. Journal of Applied Psychology, 68, 517-524.

Jackson, J. M., Buglione, S. A., & Glenwick, D. S. (1988). Major league baseball performance as a function of being traded: A drive theory analysis. Personality and Social Psychology Bulletin, 14, 46-56.

BY THE STARS

--Bill James

The genesis of this article was something I remember reading, but can't find. I read a couple of articles in baseball publications maybe a year ago which argued that there were astrological patterns in types of players. Babe Ruth was born on February 6, and Henry Aaron on February 5; both were Aquarius (Aquarians? Aquariuses? Aquariums? Whatever. . .). Tris Speaker (.344 with 117 home runs), Cap Anson (.334 with 96 home runs) and Paul Waner (.333 with 112 home runs) were all born under the sign of Aries. Power hitters tend toward the sign of Aquarius, these articles argued, while high-average hitters with line-drive power tend toward Aries; base stealers tend to be this and slick-fielding shortstops that. You get the point, I hope, because after I did the research I I couldn't find the articles which proposed this theory so I could do the setup properly. To avoid giving offense this may be just as well, but if you know exactly where these articles are I would appreciate being notified.

How do you examine such an idea? I decided to compare the question to the available methodologies. Is this a question of run creation? No. Is this a question of career projection? No. Is it a matter of defense? No, not really. Minor league value? No. Player similarity? Ahah.

It is a matter of player similarity; what these articles were arguing is that similar players have tended to be born under the same birth sign. How do you determine, using similarity scores, whether similar players have tended to be born under the same sign?

I have a BASIC data file which contains the career totals of all major league players who have played 1200 or more major league games and are not active in 1987, a total of 672 players. Actually, the data base contained only 443 players before I undertook this project, but it's 672 now. We just ask, among those 672 players, two questions:

1) Who is most similar to whom?
2) Do these players have the same birth sign?

Each of the 672 players has a "best comp" among the other 671 players--for example, the best comp for Tris Speaker is Paul Waner, the player most comparable to Charlie Gehringer is Frankie Frisch, the player most comparable to Sal Bando is Ron Cey, the player most comparable to Luis Aparicio is Rabbit Maranville, the player most comparable to Ed Bailey is Andy Seminick, the player most comparable to Dusty Baker is Gary Mathews, etc.

There being 12 birth signs, we would expect that, by random chance, there would be 56 cases in which most-similar players happened to have the same birth sign (672/12 = 56). If there is any pattern of similar players having the same birth sign, that 56 figure would shoot up quickly. The same system quickly picks up on any real similarity. Players who are 6'2" have a pronounced tendency to show as most-similar to other players who are 6'2", even though the data base has no information about any player's height. Black players have a strong tendency to show as most-similar to other black players, even though there is no code in the data base for the player's race. The system considers the defensive position to be an important element of similarity, so that a catcher almost never shows as most-similar to anyone other than a catcher, and outfielders tend to show as most-similar to other outfielders--but even if you block out those elements of the similarity formula (and base similarity strictly on hitting stats), catchers will still tend to show as most-similar to other catchers, etc. Players have a tremendously strong tendency to show as most-similar to their contemporaries. This is because things like height, race, defensive position and the era in which the player is active are real elements of the player's identity, and thus tend to seek out those same elements in looking for best-comps.

It isn't necessary under this system that players with the same birth sign share ALL of the same characteristics; it is only necessary

that, once in awhile, A FEW of the players born under the same sign would share SOME of the same characteristics. Given just that tiny edge, the number of birth-sign "matches" would rise dramatically (from 56). In all of the cases above, I am certain that the number of "matches" would exceed the random expectation by at least 100%. Guessing that there are probably about 150 black players in the data base, then there should be (by random chance) only 33 cases in which the most-similar player to a black player is another black player. In fact, there are probably more than 100 such cases. This happens because this is an EXTREME test; many players are SIMILAR to Willie Horton, but who is MOST similar to Willie Horton (George Foster, using the current system.) The importance of a small edge is tremendously magnified in an extreme test. In extreme data, even a very slight advantage will tend to manifest itself repeatedly.

I decided before I set out that the following chart would apply:

Number of "matches"	Would suggest
82 or more	Strong indication of astrological effect
72-81	Real indication of astrological effect
62-71	Faint indication of astrological effect
51-61	No indication of astrological effect
Less than 51	Bad luck

There were, in fact, 61 cases in which the two most-similar players had the same birth sign, landing on the border of "no indication" and "faint indication". I was being generous; the figure 61 is actually quite likely to result from luck.

As an experiment, I re-ran the data adding an artificial bonus of just 5 points (on a 1000-point scale) for players of the same birth sign. Given just that tiny artificial edge, the number of "matches" shot up to 77. When I created a 5-point artificial "penalty" for players having the same birth sign, the number of matches dropped to 41. Again, that demonstrates the extreme senstivity of the method to small advantages.

I approached the issue from a different angle. Taking a "random swath" of the data, I figured the average similarity score when one player was compared to another. The average similarity score for two players compared at random, based on 91,000 comparisons, was 655.1. I then travelled essentially the same random course, computing the average similarity when one player was compared to another of the same birth sign. In a study of 11,000 players with matched birth signs, the average similarity score was 654.3, a tiny bit below the random figure.

So what does the data really show? The data shows that there is, in fact, no pattern of similar players being born under the same birth sign. Birth sign is not a significant element of a player's identity.

Which, of course, most of us would have expected, anyway; I put a week into this damned thing and came out knowing the same thing that you all knew going in. But that's the nature of science: you have to check.

The beauty of baseball is that it provides such a wonderful opportunity to check. Think about it: what is this argument, but the baseball version of a universal system? The whole theory of astrological signs is that individuals with the same birth sign will tend to share characteristics, isn't it? Not that they will all show all of the same traits all of the time, but that there are patterns. It's the same in baseball as in the outside world, except that rather than being told that this group of people is passionate and that one is careful with money, we are told that these are power hitters and those are base stealers.

But in baseball, perhaps as nowhere else in the entire world, we can actually check.

No soap.

INDEPENDENT SITUATIONS AND BASEBALL STATISTICS

BY DAN HEISMAN

Runner is on 2nd, two out, bottom of the fifth in a 0-0 ballgame. Batter tenses, pitch comes in - clean single to center scoring the go-ahead run! Big hit?

Consider these four possible scenarios:

1) It is the only run of the ball game. The single is much discussed in the media the next day as a big clutch hit.

2) The go-ahead team breaks it open 9-0 and proceeds to win by 9-3. It is the Game Winning RBI (GWRBI), but as usual some fans call the local sports talk show and wonder why "anyone should get a GWRBI in a blowout, and anyhow shouldn't the GWRBI go to the guy who drove in the fourth run (which made it 4-0)?"

3) The other team ties and goes ahead in the sixth, but the home team wins in the 11th anyway, 7-6.

4) The other team goes on to win 4-1.

Lets consider the follwing factor: In terms of batter performance, did what happened afterward have any effect on the situation the batter faced in the bottom of the fifth? Was there any less pressure on the batter caused by subsequent events? Was he any less "clutch" due to which scenario occurred above?

The answers to these questions are obviously "NO". Yet when many baseball fans and even SABRMETRICIANS are analyzing the importance of RBI's, they sometimes ignore this reasoning.

For example, the Victory-Important RBI (VI-RBI) attempts to assign an importance to an RBI by the after the fact score. This ex post facto analysis does provide some additional measure, but that measure is not one of pressure, or clutch ability by the batter. A batter with a high VI-RBI measure may be no more clutch than one with a low measure, i.e. performed identically in the same situations.

The "importance" of RBI's (or any other statistic) is independent of subsequent events; i.e., they must only be measured by the situation that the batter (or whoever) is facing at the time. From the above, it is easy to conclude that any new statistic to measure batter performance should never take into account the final score.

Now I will be the first to admit that I don't like the Game Winning RBI statistic. However, if they are going to keep it as an official statistic, they are doing it right. Talk show fans are quick to point to scenario 2 above where the game is broken open and say that the batter did not deserve the game winning RBI because "this didn't cause the winning team to outscore the other team (ultimately)". This logic doesn't hold at all. Which of the

two faced the more critical situation: 1) The batter when it was 0-0 who put his team ahead, or 2) The batter who was already comforted by a 3-0 lead when he drove in the fourth run?. Was it the fault of the first batter that by the time the team was up 9-0 the pitchers were trying to avoid putting batters on base with walks and gave up three meaningless runs? In that hypothetical game only one batter faced the pressure of a tie score and came through to put his team ahead, reducing the pressure on subsequent batters. If anyone should get a GWRBI, it should be the batter that last broke the tie. And, as shown earlier, any measure of "pressure performance" must include only situations that were present at the time of the at bat, and have nothing to do with subsequent score.

Admittedly there are other factors involved. If the go ahead run scores later in the game, the batter is under more pressure than he would be in breaking a tie or driving in runs earlier. After all, the less time a team has to score and consequently allow the opponent to come back, the more critical the situation. Therefore these kind of timing factors can be taken into account. But using this same logic, the team's position in the pennant race and the time of the season could even be a larger pressure factors than the game score at the time. This is because games played where the player's team still has a legitimate shot at first place put more pressure into <u>every</u> situation than does a meaningless game later in the season between two also-rans. When their team is out of contention much pressure is removed; to many players winning itself may be meaningless and individual statistics more important. This makes game-type statistics like GWRBI relatively worthless, as only a portion of the players on either team really have a strong desire to win and consequently feel added pressure on themselves. Just witness many late September games, if you know what I mean.

In conclusion, measuring pressure performance is a tricky business. Game situation (score, inning) and team situation (standing, time of season) at the time of performance are both real factors, while the final score of the game is really meaningless.

ARE HALL OF FAME STANDARDS DECLINING?

Nothing is more common than to hear that the Hall of Fame is becoming too easy. Players like Willie Stargell, who in the old days would have had to buy a ticket to take a tour, now waltz in on the first ballot, an honor denied even to Joe DiMaggio. You know the song.

But are the Hall of Fame's standards really declining? Certainly the number of players elected on the first ballot is increasing, but that's a very peripheral honor; two years later, who is going to remember which ballot you got elected on?

First of all, I decided to check to see whether the number of people elected to the Hall of Fame was increasing. From 1969 through 1973, 30 people were inducted in the Hall of Fame. From 1974 through 1978, the number dropped to 26. From 1979 through 1983, 18 people were inducted in the Hall of Fame. And from 1984 through 1988, only 16 people were inducted.

Hm. Let's make a chart of that:

1969-73	1974-78	1979-83	1984-88
30	26	18	16

Second, I compared the Hall of Fame inductees to a set of ten basic performance standards. Those standards were:

1. A .300 batting average.
2. A .280 batting average.
3. 3000 hits.
4. 2500 hits.
5. 500 home runs.
6. 300 home runs.
7. 1500 runs scored.
8. 1500 RBI.
9. Playing a key defensive position.
10. Being a regular on three league-champion teams.

A key defensive position means catcher, second, short or center field, the "up the middle" defensive slots.

There was a similar list of standards for starting pitchers--300 wins, 250 wins, 3000 strikeouts, 2000 strikeouts, a 3.00 ERA, a 3.50 ERA, a .600 winning percentage, a .550 winning percentage, a 2-1 strikeout/walk ratio and 35 career shutouts. I made up another list of ten standards for relief pitchers, although this is only used for Wilhelm.

From 1969 through 1978, players selected for the Hall of Fame met 35% of these Hall of Fame standards.

From 1979 through 1988, players selected for the Hall of Fame met 39% of these standards--actually higher than in the previous period.

In part, this increase is illusory. Only one player in baseball history meets all ten standards, that one being Willie Mays, who entered the Hall of Fame in 1979. Along with Henry Aaron (9 of the 10), Mays dragged up the average of the era, which would be just 35% for 1979-1988 without these two; indeed, the average for the years 1984-1988 is just 30%. But at the same time, one must note that from 1969-1973, three players were selected for the Hall of Fame who met only one of the ten standards (Ralph Kiner, whose only credential was 300 home runs, Jesse Haines, whose only standard met was a .550 winning percentage, and Rube Marquard, who met only the standard of a 3.50 ERA.) Since 1979, no player has been selected who did not meet at least two of the ten standards.

In short, one cannot conclude that the standards for Hall of Fame selection are in decline. Hall of Fame errors have been made, on back to the 1940s. The BBWAA vote has been getting easier, but the Veterans Committee has been getting tougher. On balance, the Hall is holding its own.

--Bill James

RABBIT REVISITED

by Joe Mangano

In 1987, the hot topic in baseball circles was the allegedly livelier, or "rabbit" ball. Dozens of articles appeared chronicling the record-breaking number of homers flying out of the park.

Now its 1988, and the homers aren't flying as frequently; in fact, homer totals are at their lowest since 1984. Rabbit ball articles are being supplanted with theories that the 1988 ball was deadened.

Are these theories true or not? Was the 1987 ball juiced up? And was the "rabbit" taken out in 1988?

Officially, we don't know and probably never will. Denials of any changes in the ball abound, from manufacturers to Commissioner Ueberroth to league presidents. The University of Missouri-Rolla ran tests in 1987, and found the ball to be within major league standards. A lot of people remain skeptical, and continue to allege changes in the 1987 and 1988 balls. Any "proof" of an altered ball, however, can only be based on statistical inference. This article will test the hypothesis of 1987 live ball/1988 dead ball, using statistical evidence (not just home runs, but <u>all</u> relevant batting indicators).

How can this hypothesis be tested? We should first understand what happens, in theory, when the ball is changed. A presumed "rabbit" ball has an altered inner content; when propelled by another object, the ball will have greater acceleration, and hence, whether hit in the air or on the ground, will travel farther. Conversely, a deadened ball will travel shorter distances.

Even a small change in the ball can mean notable changes in statistics. Lets say, for arguments sake, that a deadened ball loses 2% of its distance. Long flies now travel six to eight feet less; therefore, home runs that previously cleared the fence by under eight feet are no longer home runs. It is not inconceivable that home runs could fall 10 or 20 percent with such a "deadened" ball - and rise by that amount using a livelier ball.

Increased acceleration and greater distance mean more than just increased home runs. Table 1 describes possible changes in outcomes when a livelier ball is hit:

Table 1

Description	Old Outcome (Regular Ball)	New Outcome (Live Ball)
1. Long fly ball outs will now fly farther, some clearing the fence.	Fly outs	Home runs
2. Long fly hits will also fly farther, some clearing the fence.	Doubles, Triples	Home runs
3. Some ground ball outs, especially hard grounders, will zip past infielders more easily.	Ground outs	Singles

4. Some soft line drives that previously fell in front of outfielders for hits, will stay in the air longer, and be caught.	Singles	Line outs
5. On hits between outfielders, the faster moving ball is harder for outfielders to cut off, and hold the batter to a single.	Singles	Doubles, Triples
6. What were line drives caught by outfielders will more easily sail over their heads.	Line outs	Doubles, Triples
7. Some infield hits (slow rollers, bunts, etc.) can now be played by infielders for outs.	Singles	Ground outs
8. What were soft liners caught by infielders are more likely to clear infielders heads, for hits.	Line outs	Singles
9. The lively ball will also accelerate faster when thrown by the pitcher, making it harder to hit.	Outs, hits	Strikeouts

Table 2 sums up how frequently each statistic occurs with the regular ball and with the live ball. A positive difference (change) suggests an increase with a livelier ball, and a negative change suggests a decrease.

Table 2

Indicator	No. of Occurences, Table 1		
	Reg. Ball	Live Ball	Change
Hits	5	6	+1
Singles	3	2	-1
Doubles	1	2	+1
Triples	1	2	+1
Home runs	0	2	+2
Strikeouts	0	1	+1

One more factor should be considered here, namely, length of the hit. The longer the hit, the greater the change using an altered ball (if a long fly travels eight feet more with a juiced ball, an infield hit may travel only one or two feet farther). Thus, the impact on home runs may be greater than the +2 indicated in the table.

We now may make rough predictions for statistical changes from 1986-87 (livelier ball) and 1987-88 (deadened ball):

Table 3

Indicator	Change, Table 2	Length of hits	Prediction 1986-1987	Prediction 1987-1988
Hits	+1	Mixed	up	down
Doubles	+1	Avg-long	up	down
Triples	+1	Long	up	down
Strikeouts	+1	-	up	down
Singles	-1	Short-avg	down	up
Home runs	+2	Very long	UP	DOWN

While there may be other factors affecting trends for these statistics, no major changes are apparent for 1986-87-88. Equipment (other than balls) is still the same. There have been no major rule changes (calling more balks in 1988 should not greatly affect the indicators mentioned). The cast of hitters and pitchers is basically unchanged (for every departing Dave Kingman, there is an incoming Mark McGwire). Batters, on average, are swinging just as hard, and pitchers are throwing just as hard. There are no new ballparks; schedules are similar; and there are no significant weather differences from year to year. Even though 1987 was the year of the corked bat for hitters and sandpaper for pitchers, it is unclear if these had any overall effect on the game; and they may well have been employed with the same frequency in 1986 and 1988.

Apparently, then, there is no major factor affecting hitting performance from 1986-87-88. The livelier ball/deader ball hypothesis may now be tested, and the prediction model applied to actual records. Table 4 displays data for the six previously mentioned indicators, for 1986, 1987, and 1988 (to date):

Table 4

				Change, 1986-87		Change, 1987-88	
Indicator	1986	1987	1988*	Pred.	Actual	Pred.	Actual
Hits/G	17.545	18.001	17.415	up	+ 2.6%	down	- 3.3%
Doubles/G	3.097	3.227	3.086	up	+ 4.2%	down	- 4.4%
Triples/G	.407	.426	.415	up	+ 4.7%	down	- 2.6%
K's/G	11.754	11.924	11.187	up	+ 1.4%	down	- 6.2%
Singles/G	12.227	12.231	12.346	down	0%	up	+ 0.9%
Homers/G	1.814	2.118	1.568	UP	+16.8%	DOWN	-26.0%

* Including games of June 6; 34% of the season

Hitting changes strongly support the predictions, although, as previously mentioned, a statistically perfect proof of the theorems is impossible. For 1986-87, the predicted large increase

in home runs came true (16.8%). The four categories for which a modest increase was predicted were clustered between 1.4% and 4.7%. Even though a predicted drop in singles didn't materialize, the figure remained constant, far below the increases of the other five.

In 1987-88 (to date), the opposite trends were forecast, and these predictions proved accurate for all categories. Home runs, predicted to drop sharply with a "deadened" ball, are down 26.0%. The four categories ticketed for small decreases are again clustered, between -2.6% and -6.2%. Singles have increased by a modest 0.9%.

These data may cause skepticism, as they are insufficient (i.e the '88 season is still incomplete), may be a statistical "coincidence" (homers have historically risen and fallen somewhat more sharply than the other indicators), or due to factors other than a rabbit ball/dead ball. Again, these criticisms cannot be scientifically quashed from these data alone. We can, however, apply these tests to other years in which ball alterations were strongly suspected.

The year 1930 may have been the greatest hitting year of all time. For example, Bill Terry's .401 broke the .400 barrier for the last time in NL history. Hack Wilson set the major league RBI record with 190. The overall batting average in the majors was an incredible .295. According to Bill James, the noted baseball statistician and historian, the 1930 ball was deliberately juiced up; and after seeing the consequences, the authorities made the 1931 ball with less "rabbit" in it.

The years 1976-77-78 are also applicable. In 1977, a fantastic number of home runs sailed out of parks. The Red Sox had four 30 homer players. George Foster hit 52 (the only player in the last quarter century to break 50). Rod Carew flirted with a .400 average, then settled for a final .388 mark. In the face of official denials, Sports Illustrated conducted a test that showed that indeed, the 1977 ball bounced higher than the 1976 model. Then the 1978 home run total dropped, and many cried that the rabbit had been taken out of the ball.

Other years also may be comparable. The 1911 season was also alleged to be a rabbit ball year, in the midst of the dead ball era. In 1920 and 1921, balls were not changed, but thrown out of play as soon as they were dirty, creating a sort of "rabbit" effect.

Each of these years are examined in Tables 5 and 6. Note that there were no major changes in the game from year to year, other than the expansion of the AL by two teams in 1977.

Table 5

"Livelier Ball" Years

Indicator	Prediction	Change 1929-30	1976-77	1910-11	1919-20
Hits/G	up	+ 3.2%	+ 4.8%	+ 9.0%	+ 7.4%
Doubles/G	up	+ 5.3%	+13.4%	+16.3%	+11.9%
Triples/G	up	+ 8.8%	+11.6%	+14.3%	+ 8.5%
K's/G	up	+11.7%	+ 6.8%	+ 1.5%	- 4.1%
Singles/G	down	+ 1.4%	- 1.3%	+ 6.4%	+ 5.8%
Homers/G	UP	+14.9%	+50.3%	+41.8%	+28.5%

Table 6

"Deadened Ball" Years

Indicator	Prediction	Change 1930-31	1977-78	1921-22
Hits/G	down	- 8.4%	- 3.5%	- 1.5%
Doubles/G	down	- 4.9%	- 3.9%	- 1.3%
Triples/G	down	-16.1%	-12.8%	- 9.0%
K's/G	down	0%	- 7.6%	- 0.9%
Singles/G	up	- 6.8%	- 0.9%	- 1.6%
Homers/G	DOWN	-31.5%	-18.9%	+12.3%

Similar to 1986-87-88, the hypotheses are generally supported by the results, accurate in five of the six categories. Home runs always had the greatest rise or fall; and the first four groups were always clustered around a smaller increase or decrease. Only singles per game showed erratic patterns of change, and generally were unsupportive of the hypotheses.

So where do these numbers leave us, in our quest to prove the ball has been changed? Again, we have no true proof, but a hypothesis strongly supported by data. In the absence of any official statements on altered balls differing from past ones, and given the inability of any technique to prove that the ball has been altered, a fairly strong, supported statement can be made that the rabbit that lived in the 1987 ball is living elsewhere in 1988.

APPENDIX

Year	Games	Hits	Singles	Doubles	Triples	Homers	K's
1910	1224	20,466	16,131	2816	1160	359	9787
1911	1218	22,192	17,093	3268	1322	509	9895
1919	1113	19,856	15,439	2922	1049	446	6870
1920	1228	23,528	18,026	3614	1258	630	7271
1921	1225	25,137	18,857	3979	1364	937	6963
1922	1231	24,891	18,646	3943	1247	1055	6935
1929	1221	24,823	17,826	4481	1167	1349	7038
1930	1232	25,848	18,242	4760	1281	1565	7932
1931	1227	23,570	16,920	4511	1071	1068	7908
1976	1939	33,291	24,850	5240	966	2235	18,745
1977	2103	37,834	26,579	6441	1170	3644	21,722
1978	2102	36,499	26,337	6186	1020	2956	20,058
1986	2102	36,880	25,702	6510	855	3813	24,706
1987	2105	37,893	25,746	6793	896	4458	25,099
1988*	706	12,295	8,716	2179	293	1107	7,898

* Through games of June 6 (34% of season)

Sources: Elias Sports Bureau (1988 data)
Reichler, The Baseball Encyclopedia (all other years)

REFLECTIONS OF A MEGALOMANIAC EDITOR

--Bill James

A few random notes on various issues:

1) Has anybody ever done a study of changes in the frequency of managerial changes over time? In the June 6, 1988 edition of Sports Illustrated, Steve Wulf wrote that "times have changed. This is the era of the One-Minute Manager and the One-Minute Coach." There followed a few stats (the average team in this decade has had 3 1/2 managerial changes) and a big chart (Job Insecurity), and the conclusion that "in simpler times, a skipper had to worry only about the performance of his team." And "maybe it is time for a change. . .in the way teams treat their managers and coaches."

What was missing in this, of course, was any attempt to establish that the change which was being lamented has actually occured. Yes, managers get fired a lot, but is this really any different than it was in 1958 (when five of the 16 managers got the axe during the season) or in 1938 or 1918? This study wouldn't be that hard to do--just count the number of men who have managed major league teams (20 or more games to eliminate interim managers) in each decade of this century, and divide that by the number of team/seasons.

I can't resist. The press holding other people to a standard which they themselves have no intention of observing is hardly news, but I've been dealing with Sports Illustrated on a now-and-then basis since 1981, and I swear they've had at least a dozen baseball editors in that time. I can't believe they've had anybody who has lasted more than a year in the job. Doesn't it seem a little bit awkward to you for an organization which changes its own manager every few months to decry the fact that baseball teams change their managers every two or three years?

2) I got a letter from Dick O'Brien, who wanted to pass along a note expanding on his earlier study of elevation as a park effect, the study which was recounted with high praise in the 1988 Abstract.

Dick's theory is that minor league team records tend to suggest that park elevation is a major--I might say even a dominant--factor in home run production. Dick compared the parks in Asheville, N.C., and Charleston, S.C., both of which were members of the Western Carolinas League in 1978 and have been in the Sally League since 1980.

By all rights, the Charleston team should hit more home runs--really, quite a lot more home runs. The dimensions in Charleston are shorter (310-402-290 as opposed to 328-404-301). The city is hotter and more humid (average temperature May-August of 78.2 as opposed to 69.3), which should boost home run output substantially more.

In spite of this, the Charleston teams over a 10-year period have averaged just 64 home runs per season, while the Asheville teams have averaged 111 home runs--73% more. The Asheville team hit substantially more home runs in each and every season.

Why? Check the elevation. The elevation of the Charleston park is 40 feet above sea level. The Asheville park is 2,140 feet above sea level.

Dick's reaction to his own info is very cautious. "I wouldn't put too much stock in the Asheville/Charleston comparison until we have comparable studies of similar cities," writes Dick--but I think he's selling his own observation short. If the park is smaller, hotter and more humid, but the teams hit dramatically fewer home runs, doesn't that focus attention rather sharply on the elevation?

3) If you haven't read the Charlie Pavitt article on the effects of being traded on performance (pages 2-3) go read it and come back.

The studies that Pavitt reports on have got to be the most phenomenal confusion of cause and effect that I have ever seen. The Jackson, Buglione and Glenwick study theorizes that players performed poorly before a trade because of the trade rumors. Obviously, what happened is not that they performed poorly because they were going to be traded, but that they were traded because they performed poorly. I mean, granting that all of us mix up cause and effect all the time, that's got to be the worst example of it on record; the trade is treated as a fixed object which creates a massive effect even before it exists!!! Excuse me, but as a general rule doesn't the cause occur before the effect?

The other study apparently makes the same blatant mistake, although in a less ghastly form, by assuming that the performance prior to a trade is indicative of the player's ability. A good analogy would be concluding that pitchers pitch poorly before they are taken out of ballgames because they are distracted (excuse me, "aroused") by the possibility of being taken out--or, in the Bateman, Karwan and Kazee study, to assuming that a pitchers performance just before being removed from a game is a standard sample of his work.

Charlie, are you sure you read that right?

4) With respect to Dan Heisman's article (go read it if you haven't, pages 6-7), several points. I developed the concept of Victory-Important RBI in the fall of 1979. Willie Stargell shared the 1979 NL MVP Award despite unimpressive performance based soley on his supposed "clutch" performance in September of that year (when, by the way, he had hit .222 for the month.) Don Baylor also won the AL MVP Award, in part also because of his clutch performance. At that time we had no access of any kind to situational statistics. There were no game-winning RBI. Attempting to deal with this complete dearth of information on a subject of such crucial import, I developed the concept of VI-RBI as a way of measuring who had driven in the runs which, as it turned out, were instrumental in winning ballgames. I never represented it as anything other than that; I always stated frankly that it was an attempt to cover as best I could an area of ignorance. Once better information on the issue became available, I immediately dropped VI-RBI; I haven't figured them or even really mentioned them, except perhaps to define the concept, for years and years. I don't know of anyone who has; the stat has been relegated, and correctly, to the scrap-heap of history.

In view of this, it seems to me that to attack VI-RBI in this way is setting up a straw man. But there is a serious issue here which Heisman is not dealing with in a serious way. The serious issue is that the apparent importance of an event appears to be radically different if perceived in a different way. Take, for example, the situation of a pennant race in which:

A) The New York Yankees beat the Detroit Tigers 4-3 on May 8th,
B) The two teams are tied for first after 161 games, when
C) The Yankees win again, 5-4.

If viewed from the perspective of the percentages AS ONE IS ABLE TO CALCULATE THEM ON THE BASIS OF KNOWN FACTS (that is, viewed as Heisman wants to view it), the second game is vastly more important than the first. The first game probably improves the Yankees' chance of winning the pennant by no more than 1%, perhaps from 14% to 15% (if that). The second game improves the Yankees' chance of winning from 50% to 100%.

But when the season is viewed as a whole (that is, when it is viewed with the eyes of God, who may be presumed to see the whole at one time), then one game is as crucial as the other, isn't it? If the Yankees lose either game, they lose the pennant. The games are of exactly equal import; the only difference was that in May, the Yankees didn't KNOW that the game was crucial.

The problem, then, is that the importance of the game is different if perceived from the viewpoint of the moment-to-moment action--what we might call the driver's-eye view--and if perceived from the viewpoint of the season as a whole (the bird's-eye view). And that is an impossible situation, because the way in which things are perceived cannot alter the reality of what they are, else the perception is false. So which is the false perception?

Well, think about it. It's obvious, isn't it? The driver's-eye-view (the one which Heisman advocates) is, of course, the distorted perception. The perception of the "driver" is distorted because the driver has unequal knowledge of the entire sequence of events. At one point he has very little knowledge about the entire sequence, and so what happens SEEMS insignificant, and at another point he has nearly-complete knowledge, so that what happens SEEMS to hold vast significance.

Let me put it to you this way: is, in fact, 50% of the pennant won on the last day of the season? Why, of course it is not. When there are 93 wins in a one-game pennant, every one of those wins is as critical as the other--and yet, looked at with the collection of facts available to the driver at that moment, it seems that way.

Why does it look that way? Because the driver at the moment of the 162 game, is UNIQUELY able to perceive the fact that every game is crucial. Every game is equally crucial--but only at one moment is the driver able to perceive the fact that it is crucial. The driver can see only a portion of the road, that which is behind him.

But, looking at the season as a whole, we can see both ways; we see his future as clearly as his past. Obviously, then, our view of the race is vastly superior to that of the driver. Isn't it? Does he know more about what happens, or do we? THE IMPRESSION THAT 50% OF THE PENNANT IS DECIDED AT THAT MOMENT IS CLEARLY A DISTORTION CREATED BY THE LIMITED KNOWLEDGE AVAILABLE TO THE DRIVER AT THAT MOMENT.

Heisman says, and I quote, that "the 'importance' of RBI's (or any other statistic) is independent of subsequent events." But that's an absurd statement, isn't it? Karl Marx, while alive, was an obscure political philosopher who wrote long, bloated tracts which nobody of importance read. Would you argue that the 20th century is powerless to change the importance of his work, that the importance of his work is fixed at the moment it is published and is independent of any subsequent event? In reality, the importance of every event in the world--and thus, of every statistic which wishes to accurately measure such event--is wholly dependent upon what happens subsequent to the event, precisely as it is dependent upon what happens prior to the event. The difference between past and future, in the eye of the driver, is that he has knowledge of one and does not have knowledge of the other. Since we have knowledge of both past and future, then to us they are all the same. Heisman wishes to say that because it is not known at the time it cannot be important--but that's obviously silly; it IS important, in every case, even though it is not known. The driver is harnassed to the moment--but we are not. To us in the future, his past and his future are all the same, all part of a continuous time line. We would be very foolish to forswear this advantage, and pretend that the only way we can see the event is as the driver sees it.

What has happened here? What has happened is that Heisman--like many other people--has confused two very distinct properties. What he is calling "importance" is not importance at all, but pressure, the pressure on the driver. PRESSURE IS CREATED BY THE PERCEPTION OF IMPORTANCE--not by importance, but by it's perception. Because the driver has the ability to see only one-half of the road (only where he has been), his perception of importance varies tremendously as he travels the road, and thus the pressure that he feels varies. But that has really nothing to do with the importance of the event.

It'll be interesting to see whether I still believe all this shit in the morning. A question: if what I am saying is correct, does it invalidate the analysis of baseball as a Markov chain? It might. What about what I was just saying about cause and effect? Doesn't treating the past and future as equal partners in a moment ignore the fact that causative waves flow only forward, not back? It might. Couldn't one oppose this argument by arguing that nothing exists in a moment except what happens at that moment? You might. If I'm ever on the Larry King show and you call in and try to make me explain this argument over the radio, will I put out a contract on your life? I might. But what I'm trying to say is that there is a serious issue here which deserves a serious consideration, and not to be dismissed as unworthy.

Two final points. The plural of RBI is RBI, not RBI's (Runs Batted In, not Run Batted In's). And for Christs sake, Heisman, learn to spell "sabermetrics".

5). Points about Mangano's article (go read in the future if you haven't in the past.)

I think that Mangano has really got the start of something here, the outline of a terrific major study, and with that in mind I hope you will tolerate me if I spend a little time talking not so much about what this article is, but about what it might be. I really think that this is a method which could contribute to our understanding of baseball in a significant way, and perhaps even to move beyond the tiny world of sabermetrics and into the broader discussion.

Many seasons have been alleged "lively ball seasons", many more than Mangano discusses. The controversy about the liveliness of the ball has jumped up in 1961, in 1947, and, really, time and time again throughout the last century.

The importance of certain seasons, then, is that they were not alleged examples, but certain ones. Mangano says that "the 1911 season was also alleged to be a rabbit ball year, in the midst of the dead ball era." In 1910 the Spalding company patented, introduced, advertised, bragged about and trumpeted incessantly it's new cork-center baseball (Cork centre base ball, as they said at the time). It was not "alleged" to be more lively; it WAS more lively. A few months later (mid-1910) the Reach Company, which made balls for the American League, introduced their own cork-center baseball, which led to increases in batting in both leagues from 1909 to 1910-1911.

The importance of this is that when we look at that period, we're not speculating; we KNOW that the ball was made more lively at that time. The known periods are a test of Mangano's thesis; if the things he says should have occured did not occur then, then something is wrong with his theory. Oh, what the hell; I'll do a little history of baseballs here.

After the brief hitting flurry--the "Cobb Spring" of 1910-1912--the business of abusing baseballs took over. The spitball had been popularized about 1905; it was still gaining momentum when the cork-center

ball came. After a couple of big-hitting years, then, came the "emery ball", the era in which players carried black stuff (soot, wet tobacco, licorice, etc.) in their gloves to deliberately discolor the ball, thus exacerbating the effects of the spitball, and negating the effects of the livelier ball.

In 1920 the ball was not changed--it absolutely, positively, unquestionably was not--but the discoloration of the ball was prohibited, and "used" (and therefore "dead") balls were thrown out of play. This made the balls livelier (again, unquestionably livelier) in 1920. As the 1920 season progressed, the pitchers complained about the balls being thrown out of play, however, at which point fate intervened. The umpires weakened a little and began to let the balls stay in play, and a player (Ray Chapman) was killed by dirty gray ball.

In 1921, then, the umpires became even more energetic about not allowing the ball to stay in play after it was dirty--and thus, the 1921 ball was even more lively than in 1920. Again, fate intervened, in the form of Babe Ruth and unprecedented popularity. The owners learned that the fans actually LIKED to see people hit. In 1925 the cork center ball was replaced by the "cushioned" cork center ball; the cork was cushioned by a little piece of black rubber, and again, home runs shot up in 1925. The ball is unquestionably responsible. The ball was juiced up again in 1930; it was deadened for 1931.

In the thirties the two leagues split; the American League was ruled by hitters, but the National League was not. In 1933 the American League ERA was 4.28; the N.L. ERA was 3.34. The A.L. remained higher year in and year out, with there being two years in which the American League ERA was a full run higher. The American League still used the Reach ball, the National League the Spalding ball. There were always rumors (reports, theories, speculation) that the balls were different, but there may well have been other causes for the discrepancy.

The United States produced no rubber. During World War II, then, the U.S. began to run out of rubber; by 1943 no rubber was available for making baseballs. The rubber was at first replaced by a substance known as balata, which is derived from a tree called the bully tree, and which was used to make chewing gum. The balata balls, however, were so dead that it was all but impossible to hit them out of the park, and (what was almost as bad), they didn't even SOUND right when you hit them; they went "splat" rather than "whack". After a couple of months the balata was replaced for the duration of the war by a hard plastic, which still didn't quite do the job but wasn't quite as bad.

By late 1946 the rubber supply was back to normal, and the balls came back to life.

From 1946 to the present, I do not know of any announced change in the composition of the baseball.

Now then, let's have an inventory. It appears to me that the contrasts to study to examine the effects of a livelier ball would be:

1911 (lively) contrasted with 1909 (dead).
1921 (lively) contrasted with 1919 (dead).
1925 (lively) contrasted with 1924 (dead).
1930 (lively) contrasted with 1931 (dead).
1947 (lively) contrasted with 1944 (dead).

The last one of those is problematical because of the personnel; there was probably a 90% turnover in personnel between 1944 and 1947. The other four could be studied, and then you could look at the war years to see if they fit the pattern.

What I would do, what I would recommend for this study, would be to start not by theorizing about what changes should take place with a lively ball, and certainly not by comparing those theoretical changes to changes which took place in a year which may or may not have had a livelier ball, but by examining carefully exactly what changes did take place in these years when we know the ball was changed. Then I would like to compare that data to Mangano's very clever chart, and then when we got to the point of looking at the 1987 season to see whether it was suspect, we would have a solid foundation to examine the season. We would KNOW exactly what we were looking for.

Historically, many of the "rabbit ball" furors have been generated by a unholy combination of simple-mindedness and paranoia. The 1961 season just roared with a lively-ball controversy, which tainted Maris' record drive. But in 1961 the American League had two new ballparks, Wrigley Field in LA and Metropolitan Stadium in Minnesota, and both were excellent home run parks (more than three home runs per game were hit in LA.) What nobody noticed at the time was that the number of home runs per game hit in the eight parks which were in the American League in 1960 actually DECLINED in 1961; the increase in home runs was wholly created by the new parks.

At the same time, it is fairly obvious that the 1988 balls cannot be precisely the same as they were in 1946; evolutionary changes have to have occurred. If you think about it, each baseball is an individual. No two balls are identical. Simple as it is, a baseball has many components--the hide (which used to be horsehide but is now cowhide), the stitches, the yarn, two kinds of rubber in the core, the cork itself. All of those things AND the way they are combined cannot be locked in time. Something like a new machine to wrap the yarn is going to have an impact. If you change the tension setting on that machine, you change the baseball--and you can't keep the tension absolutely constant on a machine while it makes a million baseballs.

So changes have certainly occurred--yet at the same time, there might well be atmospheric or weather conditions or other changes in the game which are not well understood, but which could mimic the effects of a livelier ball almost perfectly. Mangano's method is intriguing because it is a first step toward a better understanding of this omnipresent issue.