By the Numbers

The Newsletter of the Statistical Analysis Committee of the Society for American Baseball Research
Volume 2, Number 4 September, 1990

Committee News

SABR XX. The convention went very well, without last year's drama. A number of the presentations incorporated statistical methods, and we are fortunate enough to have four of these presentations for this issue of the Newsletter.

About 40 people attended the Statistical Analysis Committee meeting. We discussed what the Committee should be doing over the next year, and it looks like the bibliography project will be going forward. (I have heard from a number of you on the bibliography project, and I will writing you individually in the next week or so.)

In addition to the bibliography project, a number of people suggested finding a way of facilitating communications between people working in the same area. One idea which surfaced was for people to volunteer as "clearinghouse coordinators" for specific research areas.

Here's how this would work: Once we have volunteers to serve as coordinators in specific areas, we would hope that people working in those areas would keep in touch with the coordinators, to let them know what they are doing and what sort of progress they are making. Mark Pankin (1018 N. Cleveland St., Arlington, VA 22201, 703-524-0937) has already volunteered to serve as a coordinator for work on batting order effects (a summary of his SABR XX presentation is included in this Newsletter).

Also, although it seems like a long way off, SABR XXI will be coming up soon. Once again, I'd like to organize a session in which a series of presentations using statistical tools can be made together. If you have something you're working on and plan to be in New York for next year's convention, please let me know.

Future Issues. The December issue is looking good. I have pieces in hand from Rob Wood, Jorgen Rasmussen and Bob Davis, with something to come from John Stryker and something else I'm working on. It should be a good issue. But I always need material. A good length is 3-4 pages, typed, single-spaced (although if you send something typed, send it double-spaced). If you can send a diskette, I use Microsoft Word Version 5.0 (but any earlier version of Word will do); if you use a different word processing program, send me an ASCII file and I can convert it.

1992 and Beyond. I'm approaching the end of my second year as chair of the Statistical Analysis Committee, and about one more year is all I'll be able to keep doing this. I hope that by next year's SABR convention we can have an idea who is willing to take over the committee. The most important--and most difficult--thing is editing the newsletter and keeping in touch with the committee members (and I don't do very well at keeping in touch, as many of you know). If you are interested in taking on this responsibility, let me know and maybe we can work out a gradual transfer.

Committee Roster and Questionnaire. I hope to publish a Committee Roster, either as a part of the December issue, or separately in between. As a part of that, I have included, for Committee members, a questionnaire in this mailing. Please fill it out and return it; we need to know what you are doing, are interested in doing, and are willing to do.

Donald A. Coffin
Division of Business and Economics
Indiana University Northwest
3400 Broadway
Gary, IN 46408
219-980-6646

Bayes' Theorem: A Way To Help Us To Revise Our Opinions

By Alden Mead

Introduction. In trying to understand results in baseball (or anything else), we are frequently confronted with the necessity of revising opinions that we thought were based on sound reasoning in the light of actual events. Examples abound, but the two that we'll be talking about are the two Chicago teams in 1990.

In 1989, the Cubs won their division, the White Sox finished last in theirs. Most experts predicted that the Cubs would contend this year, and the Sox would again be also-rans. In fact, as this is written (July), the Cubs are flirting with the cellar while the Sox are contending for the lead. On the other hand, the record of the past is still there, and the experts who picked the Cubs as contenders and the Soxs as something else were (presumably) not fools. Maybe what has happened up to now is a fluke, not reflecting the true abilities of the teams. But we can't just ignore what has happened up to now. The question is, exactly how do we decide how to revise our opinions of these teams in light of the new information.

The answer is given by Bayes' theorem, a very simple result dating from the 18th century, which has recently attracted a lot of attention from statisticians. If we formulate the question correctly, it tells us exactly how to revise our opinions. To use it, we need two things: An initial estimate, expressed in terms of probabilities, and the new information, whatever it is.

The next section gives the mathematics, and can be skipped by those who don't care about such things. After that, we analyze the Cubs and Sox.

Mathematics. To use Bayes' theorem, we have to work with probabilities, which are pretty well understood by anyone who works with baseball statistics. The probability P(x) that something (called x) is true, or will happen, is a number between 0 and 1, with 0 denoting impossibility and 1 denoting certainty. In between, it's the fraction of the times that x will be true. For instance, the probability that a .300 hitter will get a hit in a randomly chosen at-bat is .300. We also need to use what are called conditional probabilities. The conditional probability P(x y) is the probability that x will happen given that y is known. Thus, if our .300 hitter is a left-handed batter that does best against righthanders, the probability of his getting a hit, given that the pitcher is a righthander, P(H r), might be .330.

If you have any experience at all working with probabilities, it's pretty easy to answer questions like "What is the probability that a .500 team will win 3 of its first 4 games?" We can express that as $P(3\text{-}1|.500)$, the probability of a 3-1 record in 4 games, given that we have a .500 team. To calculate this, we just realize that the probability of any particular sequence of three wins and one loss for such a team is $(.5)^4 = .625$, and that there are 4 such sequences (depending on which of the four games is lost), so the probability is $(4)*(.0625) = .25$. What we want, though, is to be able to change our estimate of the team based on what it's done; that is, we'd like to know such things as $P(.500|3\text{-}1)$, the probability that it's a .500 team, given that it's won 3 of its first 4 games. At first glance, it doesn't seem clear how to get at this; in fact, it's pretty easy via Bayes' theorem, which tells us how to turn around conditional probabilities--how to get $P(x|y)$ when we know $P(y|x)$, or how to get $P(.500|3\text{-}1)$ when we know $P(3\text{-}1|.500)$.

The theorem is really easy to prove. It starts from using two ways to write $P(x \cap y)$, the probability that both x and y are true:

$$P(x \cap y) = P(x|y)P(y) = P(y|x)P(x) \qquad (1)$$

All this says is that, for both x and y to be true, we must first have y given, and then x, given that y is true. Or, of course, the other way around. But, just by dividing, we get the reversal of the conditional probabilities:

$$P(x|y) = [P(y|x)P(x)]/P(y) \qquad (2)$$

That's the theorem! If you know P(x) and P(y), the total probabilities for both events, you can reverse the conditional probabilities.

To get something like P(.500 3-1), we need not only P(3-1 .500), but also [P](.500), the probability (before we get the new information) that it's a .500 team and P(3-1), the probability that it'll win 3 out of 4 games withoug knowing how good the team is. To get at this, we'll need a little more.

Let p be the true probability that our team will win a given game, so that its winning percentage over a sufficiently long season will be p. Based on all our information, we can't know p exactly, but maybe we can place it within limits. Let «p» be our best guess as to the value of p. We'll use σ to denote the "standard deviation," which means approximately that we know with 2/3 probability that the actual p lies between «p»$-\sigma$ and «p»$+\sigma$. We can conveniently represent the probability P(p) (actually the probability density, the probability per unit of change in p) by the formula:

$$[(V+D-1)!/(V-1)!(D-1)!]p^{(V-1)}(1-p)^{(D-1)}$$

where V is the number of victories, D is the number of defeats, and (X)! is "X factorial", or x(x-1)(x-2)(x-3)...(2)(1). This probability density is normalized (the total probability that p will be between 0 and 1 is just 1), and it goes to 0 at both ends, reflecting the fact that there are no invincible teams (1.000), and also no totally hopeless teams (.000).

It's a fairly easy calculation to get «p» and σ in terms of V and D, and vice-versa. The results are:

$$\text{«p»} = V/(V+D) \text{ and} \qquad (4)$$

$$\sigma = \sqrt{(VD)/(V+D)^2(V+D+1)} \qquad (5)$$

Reversing this, we can get V and D in terms of «p» and σ. This is most easily expressed in the following way:

$$V = \text{«p»}Q \qquad (6a)$$

$$D = (1-\text{«p»})Q \qquad (6b)$$

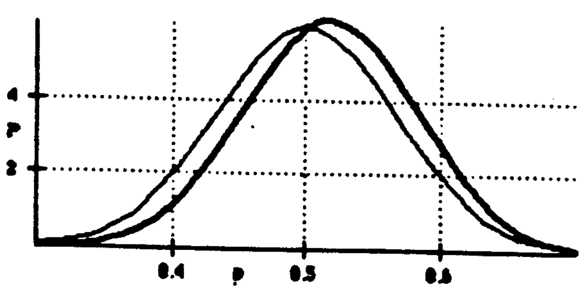
where

$$Q = [\text{«p»}(1-\text{«p»})/\sigma^2]-1 \qquad (7)$$

Now, suppose our initial estimate of a team's ability is given by values of «p» and σ, our best guess as to the prowess of the team, and an uncertainty, so that we feel that, with 2/3 probability, the team's actual ability is within σ one way or the other of «p». Suppose further that the team starts the season with v victories and d defeats. Plugging all this into the equations, Bayes' theorem tells us to revise our estimate just by replacing V by V+v and D by D+d. It's as simple as that!.

For example, suppose at the beginning all we know about a team is that it's a major league team playing in the current era. Checking all the teams from 1979-1989, and leaving out the strike season of 1981, we find an average of .500 (of course) and a standard deviation of 0.066. Using the formulas, this leads to V = D = 28.2. If now our team wins 3 out of 4, the new values would be V+v = 31.2 and D+d = 29.2 Figure 1 (below) shows how this affects our estimate. The light line shows our guess before the team won 3 out of 4. The dark line is our revised view. Note that we now view the team somewhat more favorably, but not much (the curve is shifted toward higher values of p, that is, toward higher winning percentages).

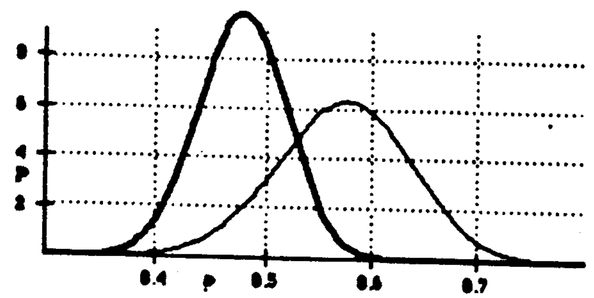
FIGURE 1



Cubs and Sox. In 1989, the Cubs had a winning percentage of .574. Suppose that, at the beginning of this year, we estimate that they have a 2/3 chance of finishing within 10 games (.062) of the same mark;

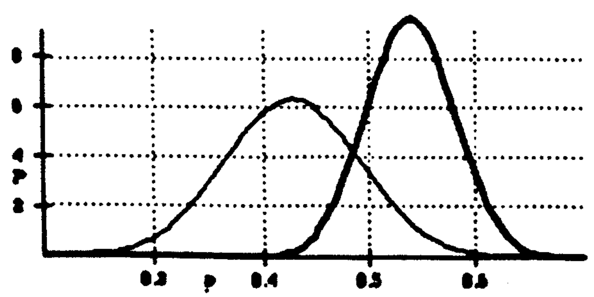
that is, «p» = .574 and σ = .062. This leads to initial values for V and D of 35.94 and 26.67 respectively. As of July 15, the 1990 Bruins had a 36-52 record, so our revised values should be V+v = 35.94+ 36 = 71.94 and D+d = 26.67 + 52 = 78.67. This leads to new values «p» = .477 (a 77-85 season record) and σ = .041 (plus or minus 6.6 wins). The graph is shown as Figure 2 (below).

FIGURE 2



As for the Sox, last year they were .429. If we start out thinking that they will finish within 10 games, or .062, of that, this leads to V = 26.91 and D = 35.82. Adding this to their 51-31 record as of July 15, we get V+v = 77.91 and D+d = 66.2=82. We then get revised estimates of «p» = .538 and σ = .041, or an 87-75 season (plus or minus 6.6 wins). Figure 3 shows us the 1990 Sox, incorporating the new information.

FIGURE 3



Defensive Average: Baseball Defense Counts Enough To Be Counted

By Pete DeCoursey

I. Defensive Performance, Like Offensive Prowess, Is Divided Into Efficiency and Power. We currently measure defense with extremely odd standards., such as visual memory and measuring plays made and errors committed, and ignoring all other balls hit at the fielder. Like batting statistics, where we measure both efficiency (BA, OBA) and power (SA), we need to measure how many of a defensive player's opportunities he turns into outs (efficiency) as well as how many outs he makes out of each opportunity, and how many bases he gives up (power).

II. Defensive Average Assigns Every Ball Turned Into An Out, Hit Or Error, and Every Ball Hit Into Fair Play To The Fielder Responsible For Playing It. Using Project Scoresheet data and field diagrams, 99% of all balls hit into play in 1988-89 were diagrammed and assigned to a fielder. Defensive average tells you how many balls were hit to a fielder's area, and what percentage he turned into outs. It also tells how many of the hits were singles, doubles and triples, which turns out to be very informative when evaluating first and third basemen. Last, DA reveals how many DP opportunities (ground balls hit to his area with a player on first and less than two outs) each infielder had, and how many DPs he started.

III. Using This Information, It Is Possible to Construct The Equivalent of Defensive Datting and Slugging Averages. By dividing the number of outs made by the number of ground balls (or in the case of outfielders, fly balls and line drives) hit to or past each fielder, you arrive at that player's Defensive Average, a kind of defensive batting average. By measuring a player's frequency of giving up extra base hits, and his ability to start DPs)for outfielders, his ability to get assists), you can measure his power, the way you would examine a player's extra base hits or stolen base columns.

It is also possible to then balance a player's offensive and defensive contributions, and compare a player's two most important ways of contributing to winning baseball games. (An Addendum to this report doing this is available from By the Numbers; please send SASE--$0.45 postage--if you are interested).

IV. Refuting the Ground Ball Fallacy. Gnarled old baseball wisdom to the contrary, hitting a ground ball through the infield is a more difficult way to reach base than hitting a ball into the air over the infield and into the domain of the outfielders. My data from 1988 and 1989 demonstrate this dramatically, as shown in Table 1 below. Remember, Defensive Average is a calculation of the number of balls hit into an area that were turned into outs. (Team-by-team data are available for an SASE--$0.25--from By the Numbers.)

		Defensive Average 1988	1989
NL,	Infield	0.737	0.723
	Outfield	0.619	0.646
AL,	Infield	0.744	0.724
	Outfield	0.609	0.611

The whole reasoning behind the platoon theory in hitting is to take advantage of a difference in batting performance which may be only 30-50 points. While it's true that errors are included in the figures above, these figures show a consistent difference of 120 points or more. And since errors put men on base, and would tend to hurt the infielders' numbers more thanthe outfielders, we see that the theory that ground balls are more likely to out a man on base than a fly or line drive is fallacious.

V. Some Data From 1989. The following charts demonstrate some critical facts about how Defensive Average illuminates defensive performance. Table 2 presents data for NL shortstops and Table 3 for AL third basemen.

A. There are Home-Road Splits in Defensive Performance. Garry Templeton finished second among NL shortstops at turning ground balls into outs, 39 points above the NL average, but finished 8th at turning grounders into outs on the road.

Table 2

NL SS	Adj DA	Adjusted DA Rankings Tot	Home	Road	DP
Smith O	.716	1	1	3	8
Larkin B	.709	2	6	1	13
Elster K	.694	3	4	2	11
Uribe J	.675	4	3	5	6
Templeton G	.670	5	2	8	9
Griffin A	.661	6	7	7	7
Thon D	.650	7	11	4	10
Owen S	.647	8	5	11	5
Dunston S	.641	9	10	7	2
Bell J	.625	10	8	12	1
Thomas A	.615	11	13	10	4
Quniones R	.604	12	9	13	14
Ramirez R	.591	13	14	9	12
Duncan M	.571	14	12	14	3

Among AL third basemen, Kevin Seitzer's home adjusted DA was .643, noticeably better than the AL average of .611. But on the road, Seitzer turned ground balls into outs at a .529 rate, 76 points below the league average of .605. Jack Howell played above the league average both at home and on the road, but still played better on the road. His road adjusted DA beat the AL average by 73 points, while his .645 home DA was 34 points above average. Carney Lansford's home DA of .596 far outshines his road average of .550 by quite a margin.

B. Defensive Ability Is A Composite Of Various Skills And Performance Levels--It Is Not Monolithic. The point here is not to denigrate the men on either list, but rather to demonstrate their performance. That is the purpose of Defensive Average: to delineate the player's record in those tasks which he must perform. To borrow a phrase, you could say that Shawon Dunston has outstanding defensive power, but needs to field for a better average. Or you could say the opposite about Ozzie Smith, Garry Templeton and Barry Larkin.

With third basemen, we can also see that each performs his three jobs at the hot corner with varying degrees of success. Those three jobs are stopping hits, stopping doubles, and starting double plays. Table 3 shows AL third basemen ranked on the first two of these tasks.

Table 3

AL 3B	Adj DA	Adjusted DA Rankings Tot	Home	Road	DP
Howell J	.661	1	4	1	6
Boggs W	.654	2	1	2	5
Gruber K	.652	3	2	4	14
Martinez E	.638	4	3		4
Gaetti G	.635	5	6	5	9
Jacoby B	.617	6	13	6	11
Buechele S	.614	7	7	8	13
Molitor P	.615	8	8	7	10
Williams E	.601	9	3		1
Martinez C	.600	10	10		2
Presley J	.593	11	9	10	12
Schu R	.584	12	15	9	16
Worthington	.583	13	12	11	7
Seitzer K	.581	14	5	13	8
Lansford C	.574	15	11	12	15
Strange D	.577	16			3
Pagliarulo M	.553	17	16	3	17

Paul Molitor clearly pursues a strategy which cuts off doubles and gives up singles (35 ground outs per double against a league average of 18, ranking first in cutting off doubles, but he ranks only 8th in cutting off hits). This kind of line-gurading strategy also appears to affect his ability to turn the double play (he ranks 11th). Wade Boggs, on the other hand, clearly emphasizes stopping hits (2nd) and turning double plays (5th) at the expense of a few extra doubles (11th--16 ground outs per double). My guess is that by stopping the extra hits he more than makes up for the doubles.

C. Errors Don't Matter--Stopping Hits, Preventing Extra-Base Hits, and Starting Double Plays Does Matter. Kelly Gruber played every day and made 22 errors. Steve Buechele, Brook Jacoby, Carney Lansford and Kevin Seitzer all made fewer errors. But they also stopped many fewer hits, and therefore were not playing as well defensively as was Gruber. Wade Boggs and Brook Jacoby both made the same number of errors, 17. Boggs stopped more hits and turned more DPs, and played measurably better defense.

VI. The Defensive Average Rules of Baseball Defense. From studying defense for the past several years, I have formulated six rules of baseball defense:

A. Defensive performance, like that of offense, is made up of efficiency and power. To measure defensive performance, we have to measure each, just as we do with statistical measurements of offense.

B. Using methods to measure the hits and extra bases a defensive player takes away, we can add them to his offensive records for the purpose of direct comparison to other players at that position, and give each player a total rating which does not ignore his defensive skills or cost.

C. Defensive players are not simply "good" or "bad" at all of the component skills which make up defensive play. Players can have extraordinary adjusted Defensive Averages and poor defensive power, or vice versa. We need to measure this so that we add some literacy to the debate over defense.

D. A player is more likely to reach base on a line drive or fly ball hit over or through the infield than by trying to hit the ball through the infield on the ground.

E. Comparison of 1988 and 1989 indicates that there can be notable fluctuations in player's Defensive Average from year to year, similar to the 20-30-40 point fluctuations commonly seen in batting averages. Therefore defense is not a constant in a player's play; good defenders, like Kelly Gruber and Chris Sabo, who were All-World in 1988, slumped in 1989.

F. Like offensive performance, defensive performance seem to have park effects, as the home-road performance splits demonstrated.

Pete DeCoursey is the Editor of the Philadelphia Baseball File, wrote the Phillies chapter of The Scouting Report '90, was associate editor of John Benson's Rotisserie Baseball Analyst, and contributes a weekly statistical feature for Jayson Stark's "Baseball Week in Review" column in the Philadelphia Inquirer.

****

Productive Efficiency and *Structural Change: Free Agency in* Major League Baseball

By Donald A. Coffin and Daniel G. Sutcliffe

I. Introduction. Firms operating in competitive markets face powerful incentives to operate efficiently; failure to operate efficiently can result in negative profits and hence failure to survive in the industry. On the other hand, firms in non-competitive markets may be able to operate at lower levels of efficiency, because they are sheltered from competition.

Such inefficiency could take any of a number of forms. For example, firms could choose to pay higher prices for inputs than would be required in a competitive setting. Or, firms could choose to operate off their production surfaces, that is, devote fewer resources to ensuring technical efficiency.

Of some interest might be a firm which sells its output in a non-competitive product market and hires labor in a monopsonistic labor market. The profits of this type of firm are sheltered both by its status as a monopoplist in its product market and by its status as a monopsonist in its product market. Such a firm (or class of firms) will, therefore, be vulnerable to reductions in market power either in the product market or in the labor market. Should the product market become more competitive, this firm will find its profits falling and may respond by attempting to reduce any technical or allocative inefficiency it has hitherto allowed itself. Alternatively, should the labor market become more competitive, this firm will again find its profits falling and again may respond by attempting to reduce inefficiency.

Major league baseball constitutes a set of firms operating in an imperfectly competitive output market and, prior to free agency, in a monopsonistic labor market. As such, a change in the structure of the labor market in major league baseball will operate to reduce team profits and engender pressure to reduce or eliminate inefficiency in operations. The key event we wish to explore in this paper is the change in the structure of the labor market for player talent, the introduction of free agency for players with six or more years of experience, which began in 1976.

We will establish that this change in labor market structure resulted in an increase in the average level of technical efficiency in major league baseball. We will also establish that the factors related to the level of efficiency with which teams operated differed in the period before players could become free agents (1962-1975) and in the later period.

In a longer version of this paper, we include sections outlining the economic theory underlying the construction of our regression model and an extended discussion of the econometric properties of this model. If you are interested in seeing this discussion, please send a 5x7 SASE, with $0.65 postage, to By the Numbers and you will receive a copy of the paper in return.

II. Estimation. The data for estimating the production function are drawn from Total Baseball. We argue in the full paper that a team's seasonal winning percentage is a good proxy for an output which is within a team's control (within limits). We must now explain our choice of player performance variables.

It will come as no surprise to readers of this newsletter that, in order to win a baseball game, a team must score more runs than its opponent. Outscoring an opponent has two components--scoring runs and preventing runs. This suggests selecting measures of player performance which are most closely associated with the desired offensive and defensive outcomes--scoring runs and preventing runs respectively. A second consideration is to keep the regression from becoming unwieldy. We decided to select one offensive indicator and one defensive indicator. A third consideration was symmetry--we wanted to use the same indicator for offense and for defense, if possible.

An anlaysis of the 1984 season indicated that a team's on-base average was the single offensive indicator most closely related to runs scored. A similar analysis indicated that the on-base average of a team's opponents, against its pitchers, was the single defensive indicator most closely associated with runs allowed. If our approach is correct, then a team's winning percentage should be closely associated with its on-base average (OBA) and with its on-base average allowed (OOBA). We found that this was correct, as the following regression for the 1962-1988 period shows (t-statistics in parentheses; $r^2 = 0.727$):

$$\ln(WPCT) = -0.631 + 2.229\ln(OBA) - 2.127\ln(OOBA)$$
$$(-6.27) \quad (30.84) \quad (-32.11)$$

It is worth noting that the coefficients on OBA and OOBA are not significantly different in absolute value and that the constant term is not significantly different from ln(0.5). This suggests that a team with identical OBA and OOBA could be expected to have a WPCT of about .500, which is convenient.

We estimated a production function incorporating team-specific production effects as

$$\ln(WPCT_{it}) = a_i + b*\ln(OBA) + c*\ln(OOBA) + v_{it},$$

where a_i are team-specific efficiency effects and v_{it} is an error term. The regression results, for the 1962-1975 period and for the 1977-1988 period, are shown in Table 1.

From the coefficients on the team-specific dummy-variables, we can construct an index of relative efficiency (with the most efficient team--the team with the largest coefficient--having a relative efficiency level of 1.00). These efficiency levels are shown in Table 2 on the following page. The mean level of efficiency in the 1962-1975 period was 0.923, with a standard deviation (σ) of 0.034. In the 1977-1988 period, the mean level of efficiency rose to 0.959, with a σ of 0.021.

In absolute terms, relative efficiency increased in the period following the abolition of the reserve clause. We tested to see whether the difference was statisitcally significant, and found that it was--a t-test for the difference in means calculated a t = 3.47, which means the difference in mean efficiency levels is significantly different from 0 at the 99% confidence level.

Table 1: Regression Results: Frontier Production Function

Variable	1962-1975 Coefficient	1977-1988 Coefficient
OBA	2.020	2.324
OOBA	-2.079	-2.170
ATL	-0.826	-0.582
CHINL	-0.745	-0.555
CIN	**-0.735**	-0.557
HOU	-0.868	-0.590
LA	-0.852	-0.577
MON	-0.821	-0.568
NYNL	-0.893	-0.597
PHIL	-0.782	-0.559
PIT	-0.756	-0.596
STL	-0.791	-0.572
SD	-0.833	-0.593
SF	-0.816	-0.582
BAL	-0.798	-0.556
BOS	-0.790	-0.570
CAL	-0.787	-0.591
CHIAL	-0.832	-0.558
CLE	-0.782	-0.634
DET	-0.771	-0.577
KC	-0.834	-0.548
MIL	-0.842	**-0.532**
MIN	-0.837	-0.574
NYAL	-0.800	-0.562
OAK	-0.830	-0.539
SEA		-0.602
TEX	-0.856	-0.597
TOR		-0.556
R^2 (Adj)	0.736	0.740
F	35.19	33.75

All coefficients are significant at the 1% level. In each regression, the "most efficient" team--the arithmetically largest coefficient--is shown in boldface.

We also tested to see whether there were significant differences in efficiency levels between teams within each time period. We did this by re-estimating our regression, excluding the dummy variable for the most efficient team, but including a constant term. If the coefficients on the remaining team variables are significantly different from 0, then these teams were significantly less efficient than the most efficient team.

Table 2: Relative Efficiency Levels		
	Relative Efficiency Level	
Team	1962-1975	1977-1988
ATL	0.913***	0.951
CHINL	0.896***	0.977
CIN	1.000	0.975
HOU	0.875***	0.944**
LA	0.890***	0.956
MON	0.918**	0.965
NYNL	0.854***	0.937**
PHIL	0.954	0.973
PIT	0.979	0.938**
STL	0.946*	0.961
SD	0.907**	0.941**
SF	0.922***	0.955
BAL	0.939**	0.976
BOS	0.946*	0.962
CAL	0.949	0.943**
CHIAL	0.908***	0.974
CLE	0.954	0.903***
DET	0.965	0.956
KC	0.906***	0.984
MIL	0.899***	1.000
MIN	0.903***	0.959
NYAL	0.937**	0.970
OAK	0.909***	0.993
SEA		0.932**
TEX	0.886***	0.937**
TOR		0.976
Mean	0.923	0.959
St. Dev.	0.034	0.021

***Efficiency level is significantly different from the most efficient team at the 1% level.
**Efficiency level is significantly different from the most efficient team at the 5% level.
*Efficiency level is significantly different from the most efficient team at the 10% level.

Note that in the 1962-1975 period, only five of 23 teams were NOT significantly less efficient than Cincinnati (Philadelphia, Pittsburgh, California, Cleveland--yes, Cleveland--, and Detroit). However, in the 1977-1988 period, only eight of 25 teams WERE significantly less efficient than Milwaukee. Not only was the average efficiency level higher, the differences between teams became less prevalent.

We should add a word about the meaning of efficiency here. What is being picked up by the team-specific variables is a contribution to winning percentage which is team-specific and not captured by the performance variables (OBA and OOBA). To put this another way, in the 1962-1975 period, Cincinnati, Philadelphia, Pittsburgh, California, Cleveland, and Detroit had consistently higher winning percentages than could be accounted for by their performance levels alone (and the Mets, for example, were worse than could be accounted for by performance alone). Note that this is not a measure of which teams had the highest winning percentages, but rather a measure of which teams used best the performances that their players produced.

III. Explaining the Changes in Efficiency Levels. The rise in efficiency levels from the 1962-1975 period to the 1977-1988 period need not be a result of free agency; other factors are also possible. For example, the addition of expansion trams could depress efficiency by adding team <u>management</u> which is less capable than the existing managements. If this were true, then the largest effect of efficiency should be found in the initial period of expansion, since it was probably less well anticipated and therefore less well prepared for.

When we examined this hypothesis, however, we found that efficiency fell from an average of 0.930 in the 1962-1968 period to 0.911 in the 1969-1975 period, then rising to 0.954 in the 1977-1988 period. The difference between the 1962-1968 and 1969-1975 efficiency levels was not significant (t = 0.97).

A second possibility was the introduction of salary arbitration in 1974. Evidence on changes in average salary levels, however, suggests that they did not begin to rise significantly, and thus did not put significant pressure on team profits, until after free agency occurred.

We concluded, therefore, that a change in the mid-1970s, plausibly connected with free agency, led teams to change their managerial behavior so as to reduce the extent of inefficiency. We then turned to a detailed analysis of efficiency levels in the two periods.

IV. Explaining Efficiency Levels. We investigated a number of measurable team characteristics to see which of these are related to team efficiency levels.

We note that more efficient teams have higher winning percentages than can be explained solely by their performance levels. However, this may be related to some other factor which varies between teams, which affects efficiency, and which we did not measure. If there is a strong relationship between efficiency levels and winning percentage, then this is a real possibility. So we examined the relationship between winning percentage (WPCT) and our measure of efficiency.

Teams experiencing greater variability in winning percentage from year to year may be experiencing performance variations or they may have less consistent, and perhaps lower, levels of efficiency. We therefore examined the relationship between variability in winning percentage (measured by the coefficient of variation--CVWPCT--in team winning percentage) and our measure of efficiency.

Gerald Scully has argued (in *The Business of Major League Baseball*) that managers experience reduced efficiency when they change teams, as they have to become familiar with a new set of players. We measured managerial turnover, both in total (MGRS), before a season (BEFORE), and during a season (DURING) and examined the relationship between this and our measure of efficiency. Extending this to teams, one can argue that teams with a high degree of player turnover may operate less efficiently. We therefore measured player turnover and examined the relationship between turnover (PLAYER) and our measure of efficiency.

We expect expansion teams to be less efficient, as they may have to employ managerial personnel with less-well-developed skills. However, this effect should be strongest for the first set of expansion teams. We examined the relationship between expansion waves (EX1, EX2, EX3) and our measure of effiency.

Finally, if team efficiency is affected by free agency, then the extent to which a team participated in free-agent player transactions should have some effect on efficiency as well. We counted the number of players signed from and lost to free agency and explored the relationship between these variables (SIGN, LOST) and our measure of efficiency.

We present simple correlations between these variables and efficiency in Table 3.

Table 3: Correlations with Efficiency

Variable	Correlation With Efficiency in 1962-75	1977-88
WPCT	0.522*	0.414*
CVWPCT	-0.451*	0.088
MGRS	-0.005	0.188
BEOFRE	0.208	0.259
DURING	-0.192	0.071
PLAYERS	-0.337	-0.246
EX1	-0.426*	-0.381*
EX2	-0.207	0.274
EX3		-0.069
FA		-0.008
SIGN		-0.176
LOST		0.208

*Significant at the 5% level.

Teams with higher winning percentages are more efficient, so there may be an excluded variable problem. In 1962-75, teams with more variation in winning percentages were less efficient, as were the first expansion teams. In the second period, the first expansion teams continued to be less efficient.

Correlations between player turnover and efficiency were almost significant (and negative), so more player turnover does appear to have a deleterious effect on efficiency. A regression analysis of efficiency levels in the 1977-1988 period suggests that teams that signed more free agents were significantly less efficient, while teams that lost more free agents were significantly more efficient.

V. Conclusions. We believe we have established that relative efficiency levels increased in the mid-1970s and that this increase is related to the creation of a free agent labor market for player talent. We have further established that participation in this free agent labor market did have an effect on team efficiency levels. We have not explored the specific managerial responses of teams to free agency, and that remains a promising area of research. We note, however, that managerial turnover increased in the 1977-1988 period (in total, before, and during seasons). Player turnover also increased. Neither of these increases was statistically significant, however. We have no observations on player development activities or on front-office marketing and promotional activities, both of which may have changed in response to free agency. Much work remains to be done on these factors.

Donald A. Coffin is Associate Professor of Economics at Indiana University Northwest; Daniel G. Sutcliffe is an actuary for Country Companies Insurance Companies in Normal, Illinois.

****

Batting Order Analysis Using Markov Chain Models

By Mark D. Pankin

Scoring runs for a baseball team involves "stringing together" events for which the long-term, or "expected" probabilities may be observable. This makes use of Markov chain models highly suitable for analyzing the effects of a particular batting order, or of changes in that batting order, on the numberof runs a team can be expected to score.

A Markov chin model is based on transition probabilities in specific runner-out situations and creates a matrix of such probabilities. For example, if, for the lead-off hitter in a game there are five possibilities (make an out, reach first, reach second, reach third, reach home), and if we know the probability for each of these five events, then the second hitter faces each of these possibilities with a known probability. Given what the second hitter can do (and there are more than five possibilities here), we then have a matrix of probabilities following two hitters, as shown in Table 1 at the top of the following column.

The third hitter then faces all of the <u>combinations</u> of the events for hitters one and two, again with a known set of probabilities. This permits us to create a probability matrix for hitter 3, hitter 4, and so on.

Table 1

Hitter Two Events	Hitter One Events 1	2	3	4	5
1	P_{11}	P_{12}	P_{13}	P_{14}	P_{15}
2	P_{21}	P_{22}	P_{23}	P_{24}	P_{25}
3	P_{31}	P_{32}	P_{33}	P_{34}	P_{35}
4	P_{41}	P_{42}	P_{43}	P_{44}	P_{45}
5	P_{51}	P_{52}	P_{53}	P_{54}	P_{55}
6	P_{61}	P_{62}	P_{63}	P_{64}	P_{65}
7	P_{71}	P_{72}	P_{73}	P_{74}	P_{75}

Once all of the transition matrices have been created, one can then "run" the model. The model "re-sets" every time three "outs" have been made (so that the next hitter then becomes hitter one), and then the transitions re-commence. This model can then calculate expected (long-run average) runs per inning and runs per game, based on the transition probability matrices and on the lineup used. Note that the transition matrices will vary depending on the lineup used, since different players have different performance probabilities in different situations.

The model can be used to test theories about lineup construction. It can also be used to gain insights, such as where in the lineup to bat certain players, or how to maximize the runs scored in the first inning, or to find the "best" lineup for a given team.

This study reports on research undertaken on lineup construction for the 1986 Boston Red Sox. The use of an AL team simplifies the analysis, because of the presence of the designated hitter. Boston simplifies the analysis because John McNamara fielded virtually a set lineup all year and because it was a very slow team, unlikely to steal a base. Finally, McNamara made one significant lineup switch during the year, switching Wade Boggs from lead-off to second in the order and Marty Barrett from second to first.

The study used player data for each runner/out situation from Project Scoresheet and Elias scoring position statistics for less common situations. Team averages were used for stolen bases, caught stealing, wild pitches, passed balls, and other unusual situations. Sacrifice bunts attempted and successful were excluded.

McNamara used an extremely stable lineup at eight positions, and used three players at shortstop (whose statistics were combined to create a composite shortstop). He also made only one significant change in his lineup order, switching Wade Boggs from second to leadoff on August 6 (and moving Marty Barrett from leadoff to second).

In addition to examining McNamara's lineup choices, the study presents two other explicit lineups, based on a theory of lineup construction in the 1988 Elias Baseball Analyst. This theory proposes five guidelines to be used in constructing a lineup:

1) Cluster players in the lineup by slugging average and home run percentage.
2) Spread out the players with the highest batting averages.
3) Lead off with the player with the highest OBA.
4) Bat the best contact hitter second.
4) Bat the hitter with the best home run percentage lower than fourth.

This allows us to examine explicitly four lineups, as shown in Table 2.

Table 2

McNamara's Two		Two Based on Elias	
Lineup 1	Lineup 2	Lineup 3	Lineup 4
Barrett	Boggs	Boggs	Boggs
Boggs	Barrett	Buckner	Buckner
Buckner	Buckner	Evans	Evans
Rice	Rice	Rice	Rice
Baylor	Baylor	Baylor	Baylor
Evans	Evans	Gedman	Barrett
Armas	Armas	Barrett	Gedman
Gedman	Gedman	Armas	Armas
SS	SS	SS	SS

Applying the Markov chain model to these four lineups yields the predicted inning-by-inning and average per game runs scored shown in Table 3. While lineup construction does make some difference, the differences are not terribly large. The largest difference (between lineup 2 and lineup 4) is 0.133 runs per game (2.4%), or about 21 runs per year--enough to win about two more games.

Table 3

	Lineup			
Inning	1	2	3	4
1	.714	.655	.760	.770
2	.575	.583	.556	.561
3	.615	.618	.639	.637
4	.601	.598	.589	.588
5	.600	.604	.612	.612
6	.606	.603	.603	.602
7	.599	.602	.606	.605
8	.605	.604	.606	.606
9	.600	.602	.605	.604
Total	5.519	5.473	5.576	5.586

Lineups 1 and 2 are actual lineups used by the BoSox, so we know how well they actually did with those--they scored 4.55 runs per game with lineup 1 (before August 6) and 5.64 per game with lineup 2. However, Buckner, Evans, Boggs, and Gedman all hit much better in August and September than they had earlier, while the other batters hit about the same. When 44% of your lineup starts hitting better and no one noticeably tails off, you should score more runs. What we do not know is whether the improvement was a result of the change in the lineup, although it seems unlikely that a small lineup change would lead to such a large change in offensive production.

It is also possible to create lineups randomly and examine run scoring to find each player's best lineup position, to determine runs/inning profiles of the highest scoring lineups, and to get ideas useful for building good lineups. With nine players and nine positions, there are 40,320 possible lineups. Nine lineup groups were created--600 for each player leading off. Table 4 shows the percentage of the maximal lineups with each player in each position.

Table 4

Player	Batting Order Position 1	2	3	4	5	6	7	8	9
Barrett	16	5	4	10	18	13	13	9	12
Boggs	27	27	1	21	8	6	4	4	2
Buckner	4	8	11	5	11	15	15	17	16
Rice	11	11	9	10	14	12	13	11	12
Baylor	12	6	20	9	7	12	12	11	11
Evans	8	1	49	9	4	9	10	9	4
Armas	7	4	1	21	22	11	8	13	13
Gedman	8	12	6	12	10	14	15	12	13
SS	7	27	0	5	7	10	11	15	18

The table is to be interpreted as follows: In 16% of the maximal lineups, Marty Barrett batted first; in 5% of the maximal lineups, he batted second; and so on (each row should add to 100%, except for rounding error). The idea about lineup construction is, then, to try to bat each hitter in the position in which he hits in the largest percentage of the maximal lineups. This leads to the following suggestions:

1) Evans should bat third.
2) Boggs can bat leadoff, second, or fourth.
3) The composite shortstop makes a remarkably good #2 hitter.
4) Armas should bat fourth or fifth.
5) Buckner should bat sixth or lower.
6) Rice can bat almost anywhere. in the order.
7) Barrett should bat leadoff or fifth.
8) Gedman should probably bat sixth or seventh.
9) Baylor is hard to place, if Evans bats third (but based on all of this, seventh or eighth looks about right).

The best random (and overall) lineup found in the search was: Rice, SS, Evans, Boggs, Armas, Barrett, Gedman, Baylor, Buckner. It scored 5.7 runs per game. The worst random lineup found was: SS, Gedman, Armas, Buckner, Evans, Barrett, Rice, Boggs, Baylor. It scored 5.19 runs per game.

An analysis of runs scored per inning in the 5400 different lineups generated was undertaken. Almost all of the differences in runs per game comes from differences in the first inning; lower-scoring lineups tended to score more runs in the second inning than did higher scoring lineups. About 2% of the lineups scored more than 5.6 runs per game (the top group), while 22% of the lineups scored between 5.45 and 5.5 runs per game.

What conclusions can we reach? Four conclusions seem to come out of the study:

1) Batting order differences are important. The range from best to worst was about 0.5 runs per game, or 80 runs per season, enough to win about an additional eight games. McNamara's lineups were decent, but the lineups based on the Elias suggestions were about 10 runs per year better (one win), while the overall best lineup found was about 29 runs per season (three wins) better.
2) The Elias suggestions seem to be productive.
3) The best lineups seem to maximize run scoring in the first inning. In several places, Bill James has raised the issue as to whether maximizing scoring in the first inning is productive; these results seem to suggest it is.
4) Markov chain analysis is a useful tool. Although it makes the assumption that the transition probabili-

ties are know a priori, we do have substantial data about player performance in specific situations which we can use. It is also a much faster method than most simulation methods.

Some limitations of the current results are also important. First, the results are based on the analysis of a single team for a single year. Testing these results for more teams and more years will add to our knowledge. Second, there is a question whether the Markov model can be adapted to a team with a Ricky Henderson , or to a team which employes base stealing and certain fast runners as a significant part of its offense. If it cannot, simulation methods would seem to be the only alternative. The author's feeling is that the Markov model can be adapted; he suggests modifying the transition probability matrices for given players based on the lineups being analyzed. The idea is to determine who is likely to be on base (or more precisely the probabilities of specific runners being on in specific situations), and then adjust the transition probabilities appropriately based on available data. This method requires additional steps before analyzing the lineups, but does not significantly increase the computational requirements. Since the primary goal of lineup analysis is to compare lineups, these adjustments do not have to be precise.

Third, there is a problem with state-dependent performance. The author writes "I can almost hear baseball 'insiders' saying things like 'everyone knows Evans can't hit in the third spot,' or 'if you don't have someone like Baylor behind Rice, he'll never get a pitch to hit.' I have two answers to that. First, if you think you know all the answers, you'll never learn anything and never do any better. (Some of the decisions made by McNamara and Dave Johnson in game 6 of the 1986 World Series make you wonder if they know anything at all; and responses such as sticking with the players who got me here--even if they can't bend over to field a ground ball--reinforce that impression.) Second, the assumptions underlying the model can be changed, and we can see if the results change."

Mark Pankin works in operations research. This report on his presentation was prepared by the editor of By the Numbers; he believes it is an accurate account, but welcomes corrections or additions.

****