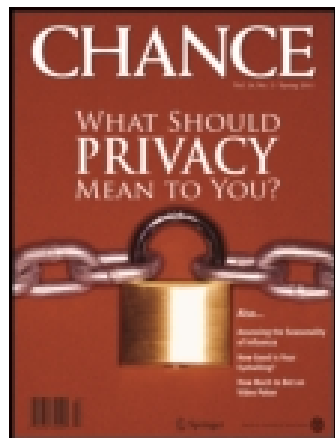


This article was downloaded by: [University of Dallas Blakley Library]

On: 26 December 2014, At: 09:45

Publisher: Taylor & Francis

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



CHANCE

Publication details, including instructions for authors and subscription information:
<http://www.tandfonline.com/loi/ucha20>

Career Trajectories in Baseball

Teddy Schall & Gary Smith

Published online: 20 Sep 2012.

To cite this article: Teddy Schall & Gary Smith (2000) Career Trajectories in Baseball, CHANCE, 13:4, 35-38

To link to this article: <http://dx.doi.org/10.1080/09332480.2000.10542233>

PLEASE SCROLL DOWN FOR ARTICLE

Taylor & Francis makes every effort to ensure the accuracy of all the information (the "Content") contained in the publications on our platform. However, Taylor & Francis, our agents, and our licensors make no representations or warranties whatsoever as to the accuracy, completeness, or suitability for any purpose of the Content. Any opinions and views expressed in this publication are the opinions and views of the authors, and are not the views of or endorsed by Taylor & Francis. The accuracy of the Content should not be relied upon and should be independently verified with primary sources of information. Taylor and Francis shall not be liable for any losses, actions, claims, proceedings, demands, costs, expenses, damages, and other liabilities whatsoever or howsoever caused arising directly or indirectly in connection with, in relation to or arising out of the use of the Content.

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden. Terms & Conditions of access and use can be found at <http://www.tandfonline.com/page/terms-and-conditions>

In this era of expensive long-term contracts, the authors' analysis of typical career trajectories is must reading.

Career Trajectories in Baseball

Teddy Schall and Gary Smith

MAJOR LEAGUE BASEBALL rookies are scrutinized carefully for evidence of their ability to succeed in the big leagues. Many are weeded out. Others have long careers. We investigate here the typical performance record for baseball hitters and pitchers during the course of their careers. We also consider whether the length of a player's career can be predicted accurately from his first-year performance. When a team offers a promising young player a long-term contract, is it buying a star or rolling dice? It turns out that first-year performance is an important predictor but that there is a great deal of individual variation.

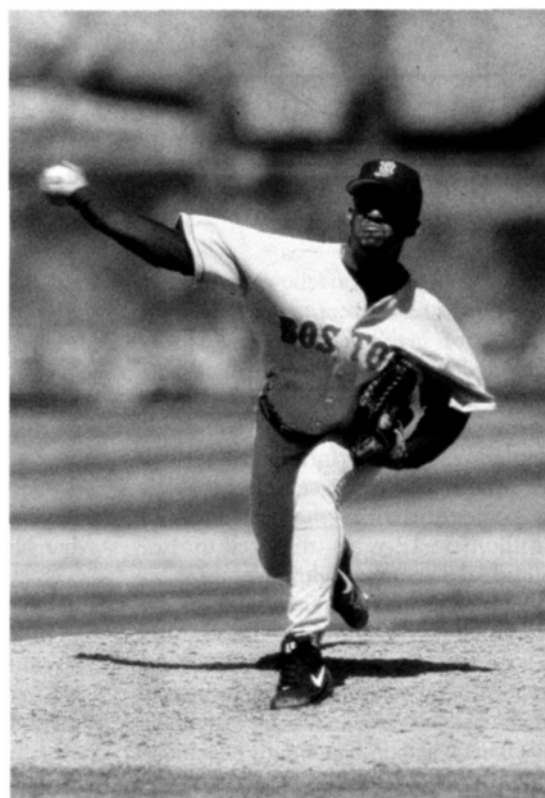
Data

We examined the season batting averages (BA's) and earned-run averages (ERA's) for all major league baseball players from 1901 through 1999, using the data in the sixth edition of *Total Baseball* (Thorn, Gershman, Palmer, and Pietrusza 1999), which is available on the Web site www.baseball.com. To reduce the influence of outliers on our performance measures, we excluded any season in which a player had fewer than 50 official times at bat or 25 innings pitched. To study career lengths, we further restricted our attention to persons who did not play in the 1999 season, presumably because they had retired. A few players sat out the 1999 season because of injuries or were temporarily relegated to the minors and will

resume their careers at a later date. The exclusion of these few exceptions, however, seemed preferable to restricting our analysis to persons who had not played in, say, the last five years.

Our final database contained 4,728 hitters, with careers ranging from 1 to 24 seasons, and 3,803 pitchers, with careers ranging from 1 to 26 seasons. The mean career length is 5.6 years for hitters and 4.8 years for pitchers. To compare performances in different seasons, each player's performance was standardized by subtracting the mean for all players that year and dividing this difference by the standard deviation of performance across players that year. We reversed the signs of the standardized ERAs so that, like BAs, high standardized values represent above-average performance.

Figure 1 shows the standardized season batting averages of three extraordinary ballplayers—Ty Cobb, Willie Mays, and Bob Uecker. Cobb's standardized BA was $-.173$ his rookie year; this was followed by an unparalleled run of 23 seasons in which his lowest standardized BA was 1.018 in his last season. Mays batted $.465$ standard deviations above the major league aver-



Pedro Martinez was recognized early on as a top prospect; now he is the dominant pitcher in the American League.

age his first year and $.389$ standard deviations below his second year. After many years of outstanding performances, his standardized BA fell to $-.963$ in his last season. Uecker is noteworthy for being allowed to play five below-average years; his worst year was his last, when he batted 2.017 standard deviations below the major league average. Uecker later achieved fame as a baseball announcer

G. NEWMAN LOWRANCE • THE SPORTING IMAGE PHOTOGRAPHY

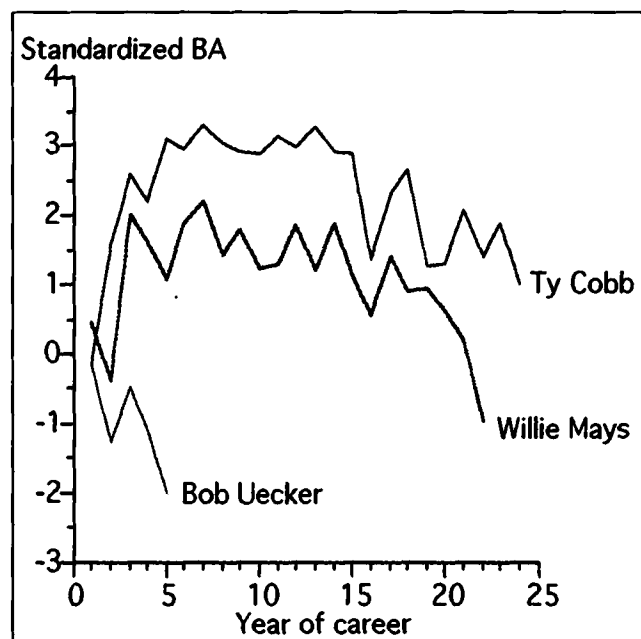


Figure 1. Three remarkable hitters.

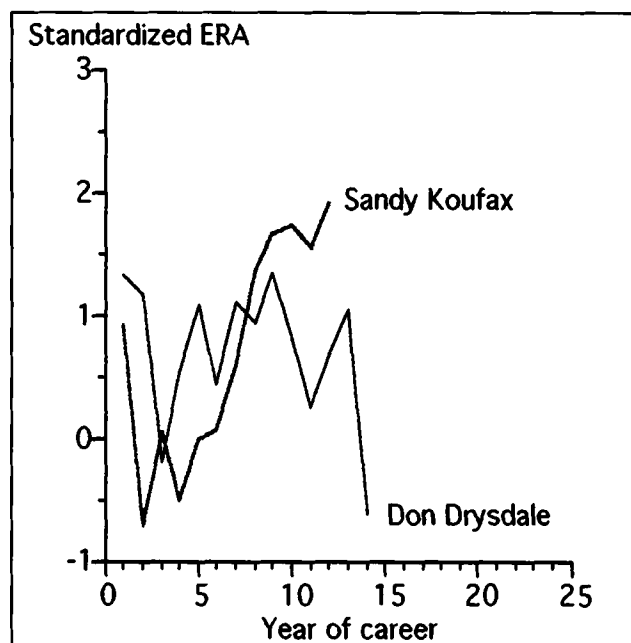


Figure 2. Two pitching teammates.

and personality, often telling stories from his "illustrious" career.

Figure 2 shows the standardized ERAs of two Dodger teammates, Don Drysdale and Sandy Koufax. As is true of most pitchers with relatively long careers, each outperformed established players in his rookie season. Drysdale was a better-than-average pitcher for 11 of the next 12 years, then fell to .627 standard deviations below average in his final year. Koufax is a remarkable exception to the typical pitching career. After a good rookie year, he had five mediocre years, followed by six very, very good years. His best season was his last, when his ERA was 1.925 standard deviations better than the major league average.

Survival Probabilities

Figure 3 shows the yearly survival rates for baseball players — that is, the probability of playing at least one more year if a player has played a given number of years. (The 32 hitters and 31 pitchers with careers of 20 years or longer were consolidated, with the reported fraction at year 20 showing the average survival rate for all of the remaining years.) The survival rate is below average in the first year (77% for hitters, 73% for pitchers), strong for the next several years (peaking at 87% in year 7 for hitters, and at

84% in year 5 for pitchers), and then generally declines. The relatively low survival rate in the first year is perhaps because the additional information gained about a player's skills during his first year playing against major league competition is much greater than the additional information gained during, say, a player's fifth season. The somewhat higher survival rates for the next several years for those who make it past the first year may be because their skills are improving during this period. We will return to this conjecture.

A logit model was used to predict the probability P of playing another season in the major leagues as a function of the number of seasons t that a person has played in the major leagues, the player's current season performance Z , and the average performance over the player's career (\bar{Z}):

$$\ln \left[\frac{P}{1-P} \right] = \beta_0 + \beta_1 t + \beta_2 Z + \beta_3 \bar{Z} + \epsilon$$

The estimates of the coefficients (and their standard errors) are shown in Table 1. These estimates are derived using the method of maximum likelihood, assuming that the multiple observations for a single player are independent. The effects of seasons are small but highly statistically significant. The effects of current performance are

substantial and statistically compelling; the effects of career performance are much smaller and, for hitters, not statistically significant.

Evaluated at the mean values of the explanatory variables, the estimated coefficients imply that each additional year in the major leagues reduces a player's survival probability by .008 for hitters and .005 for pitchers. A one-standard-deviation improvement in current performance increases the survival probability by .103 for hitters and .102 for pitchers; a one-standard-deviation improvement in career performance increases a player's survival probability by .003 for hitters and .015 for pitchers.

Career Performance Patterns

Figure 4 shows the mean value of the standardized BA and ERA in each year of the careers of the players in our database, again consolidating players with careers of 20 years or longer. Players typically perform below the league average in their first year in the major leagues — on average, .337 standard deviations below for hitters and .176 standard deviations below for pitchers. The fact that the average standardized first-year ERA is better than the average standardized

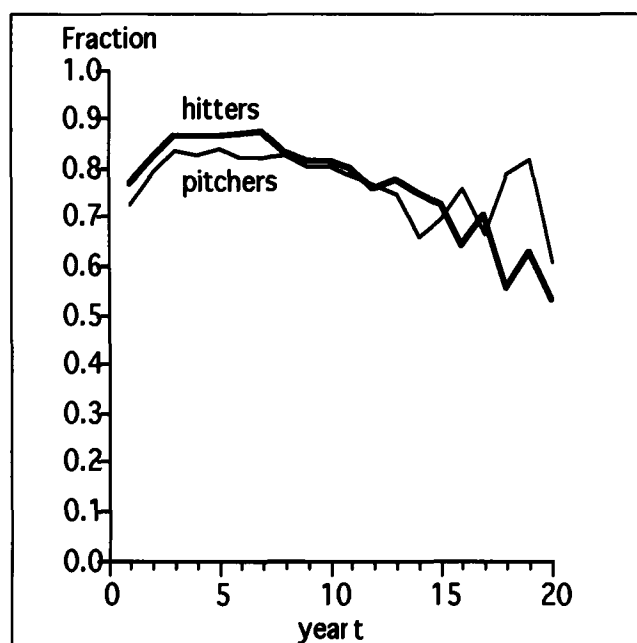


Figure 3. Fraction of those playing t years who play at least one additional year.

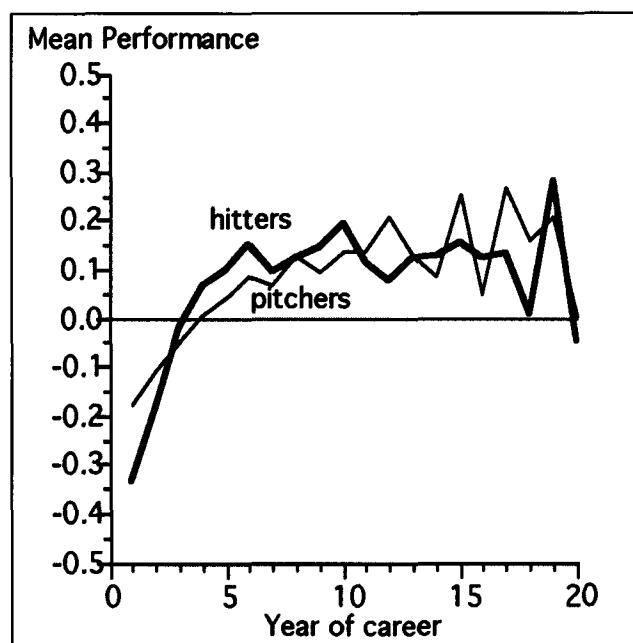


Figure 4. Mean standardized performance, by year of career.

first-year BA means that, in comparison to established players, pitchers tend to do better in their rookie seasons than do hitters. Perhaps it is easier to predict a pitcher's initial success in the major leagues. Or a rookie pitcher's style may present more of a challenge to opposing hitters than does a rookie hitter's style to opposing pitchers.

Figure 4 shows that the mean performance increases after the first year, though this may be due more to weak players being weeded out than to players improving with experience. Remember, 23% of the rookie hitters and 27% of the rookie pitchers don't play a second year. To disentangle the confounding effects of player skills changing during their careers and career lengths depending on player skills, we grouped the players by career length and examined the average career trajectory for

the players in each career group. Other influences surely matter too — a player's health, teammates, ballpark, and skills that are not measured by BA or ERA. In reasonably sized samples, these idiosyncratic factors will average out and allow us to see the systematic patterns. We consequently restricted our analysis to career lengths for which we have data on at least 100 players — up to 13 years for batting averages, 10 years for earned run averages. Figures 5 and 6 show the results.

First, notice that for hitters and pitchers of all career lengths the average standardized terminal-year performance is substantially below 0, the average performance of all major league players. The best average terminal-year values are $-.502$ for ERAs and $-.539$ for BAs, both for players terminated in their tenth year. Termination after such poor per-

formances is not surprising because the decision to terminate is based on perceived inability to perform. Remember our earlier finding that the average first-year performance is $-.337$ for hitters and $-.176$ for pitchers. It makes sense to replace an older player with a standardized performance of $-.5$ with a rookie whose expected standardized performance is $-.3$ or $-.2$, particularly if the older player is more expensive. The precipitous decline in performance near career end challenges the wisdom of long-term contracts and confirms the wisdom of Branch Rickey's dictum that it is better to trade a veteran a few years early than one year too late.

Second, hitters who last more than a few seasons tend to improve each year until the decline that precedes their termination. Pitchers, in contrast, typically do not improve with experience; pitchers with long careers just seem more successful in postponing the deteriorating performance that will end their careers. Perhaps hitters improve as they become accustomed to major league pitchers, whereas pitchers wear out their arms.

It should be emphasized that the patterns in Figs. 5 and 6 do not say that every player with a standardized performance of below $-.5$ should be terminated. There is considerable variability in performance from year to year as we see in the next section.

Table 1 — Estimated Coefficients and Standard Errors for Logit Model Predicting Survival Probability

	Batting averages		Earned run averages	
	b_i	SE	b_i	SE
Constant	2.010	.030	1.625	.031
Seasons	-.058	.004	-.031	.005
Performance	.794	.025	.677	.026
Average performance	.022	.033	.100	.036

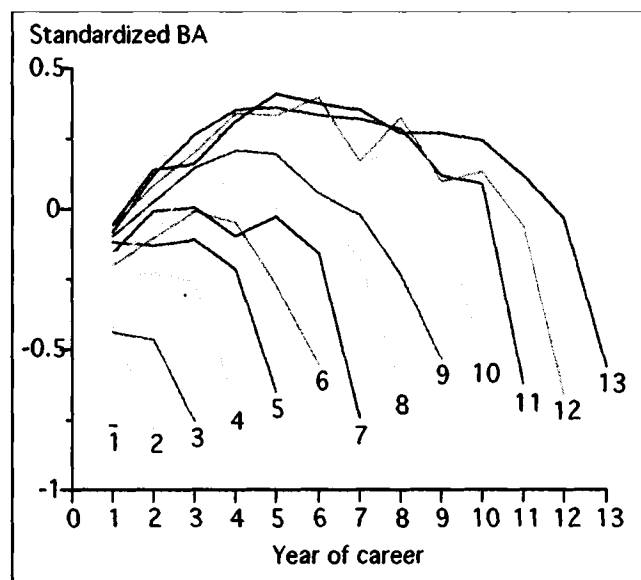


Figure 5. Batting average patterns over a career, by career length.

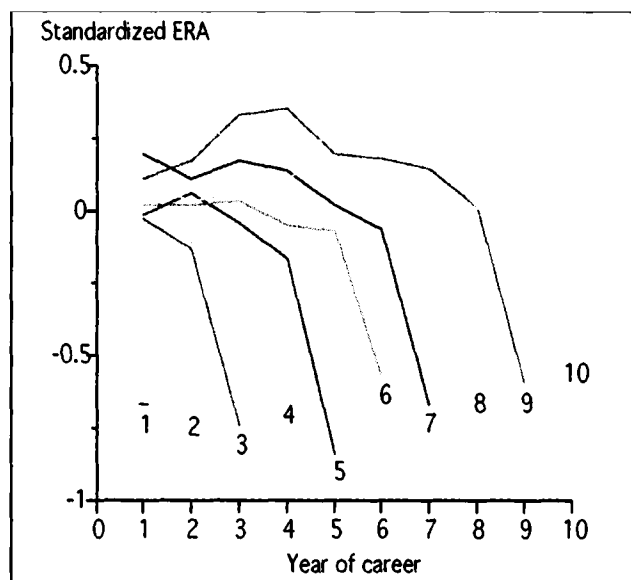


Figure 6. Earned-run average patterns over a career, by career length.

First-Year Performance and the Length of a Player's Career

For both hitters and pitchers, there seems to be a positive relationship between first-year performance and the length of a player's career. We examine this more closely. Figure 7

shows a scatterplot of career length versus first-year performance for batting averages; the pattern for earned-run averages is similar. Although these data decisively reject the null hypothesis that there is no relationship between career length and first-year performance (the correlation is .23 for batting averages and .21 for earned-run averages), the relationship is far

from perfect — explaining only 5% of the variation in career length. The estimated standard error of prediction for a given first-year performance is approximately four years.

Conclusions

Part of the charm of baseball is that there are clear tendencies that are logical and statistically persuasive, even while specific outcomes deviate substantially from these tendencies. Baseball seasons have far more games than other sports, allowing the law of large numbers to work its magic and allowing plentiful opportunities for short-term departures from long-run tendencies. A batter can go 0 for 4 one game and 4 for 4 the next. A pitcher can throw a shutout one game and give up five earned runs in the first inning of the next game. A hitter can bat .350 one season and .250 the next; a pitcher can have a 1.50 earned-run average one season and 5.00 the next. A team can win 100 games in a season and lose a World Series in four straight games. So it is with baseball careers—logical, statistically persuasive tendencies from which individual players can deviate greatly.

Reference and Further Reading

Thorn, J., Gershman, M., Palmer, P., and Pietrusza, D. (editors) (1999), *Total Baseball: The Official Encyclopedia of Major League Baseball* (6th ed.), New York: Total Sports.

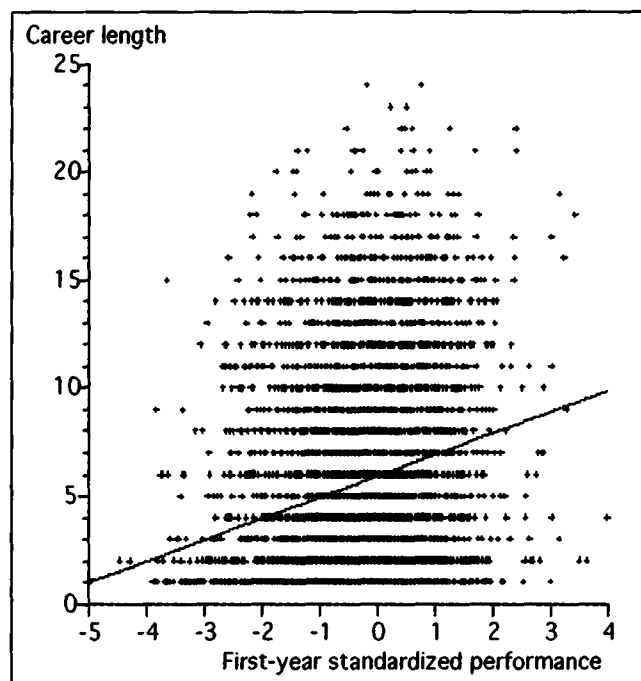


Figure 7. Batting averages, predicting career length from first-year performance.

from perfect — explaining only 5% of the variation in career length. The estimated standard error of prediction for a given first-year performance is approximately four years. An individual player's career cannot be predicted confidently from his first-year batting average or earned-run average.

The reason is not hard to find. The correlation between player performance from one season to the next is highly statistically significant, but the correlation coefficients