A Statistical Look at Roger Clemens' Pitching Career

Eric T. Bradlow, Shane T. Jensen, Justin Wolfers, and Abraham J. Wyner

Baseball is America's pastime, and with attendance and interest at an all-time high, it is clear baseball is a big business. Furthermore, many of the sport's hallowed records (the yearly home run record, the total home run record, the 500 home run club, etc.) are being assailed and passed at a pace never before seen. Yet, due to the admitted use and accusations documented in the "Mitchell Report" of performance-enhancing substances (PESs), the 'shadow' over these accomplishments is receiving as much press, if not more, than the breaking of the records.

A particularly salient example comes from a recently released report by Hendricks Sports Management, LP, which led to widespread national coverage. Using well-established baseball statistics, including ERA (number of earned runs allowed per nine innings pitched) and K-rate (strikeout rate per nine innings pitched), the report compares Roger Clemens' career to those of other great power pitchers of his era (i.e., Randy Johnson, Nolan Ryan, and Curt Schilling) and proclaims that Roger Clemens' career trajectory on these measures is not atypical. Based on this finding, the report suggests the pitching data are not an indictment (nor do they provide proof) of Clemens' guilt; in fact, they suggest the opposite.

While we concur with the Hendricks report that a statistical analysis of Clemens' career can provide *prima facie* 'evidence' (and a valuable lens with which to look at the issue), our approach provides a new look at his career pitching trajectory using a broader set of measures, as well as a broader comparison set of pitchers. This is important, as there has been a lot of recent research as to what are the most reliable and stable measures of pitching performance. Our attempt is to be inclusive in this regard.

Even more important, one of the pitfalls all analyses of extraordinary events (the immense success of Clemens as a pitcher) have is "right-tail self-selection." If one compares extraordinary players only to other extraordinary players, and selects that set of comparison players based on their behavior on that extraordinary dimension, then one does not obtain a representative (appropriate) comparison set. By focusing on only pitchers who pitched effectively into their mid-40s, the Hendricks report minimized the possibility that Clemens would look atypical.

Here, we use more reasonable criteria for pitchers that are based on their longevity and the number of innings pitched in their career to form the comparison set, rather than performance at any specific point. Thus, the focus of this paper is an analysis of Clemens' career using a more sophisticated and comprehensive database and, based on that, what one can say about Clemens' career.

A Closer Look

Before we begin our full analysis and discussion, we first take a closer look at Clemens' career. To be sure, this unavoidable act of data 'snooping' was part of our research method, and it is instructive to unfold our insights in the order in which they actually occurred. For the average fan, the most salient measures of success are winning percentage and ERA, which are a good place to start. Of course, for each game, there is a winning pitcher and a losing pitcher (hence 0.5 is the average winning percentage), and an average ERA varies between 4.00 and 5.00, which has been fairly stable over the last 30 years or so. In this light, one can see in Figure 1 how extraordinary Clemens has been in various stages of his career.

In particular, what this figure shows is that Clemens quickly established himself as a star and, in the early 1990s, he lost his 'relative' luster. His final four years with the Red Sox were certifiably mediocre (compared to his history), so much so that the future Hall of Famer was considered to be in the "twilight of his career." However, as our graph clearly demonstrates, Clemens recovered and climbed to new heights at the comparatively old age of 35. His last few years showed a second period of decline.

Now, any well-read student of baseball understands that winning percentage and ERA are fairly noisy measures of quality. Both are readily affected by factors outside a pitcher's ability, such as fielding and the order in which batting events occur. Additionally, winning percentage critically depends on run support. Analysts who specialize in pitching evaluation use measures of component events instead, such as rates of strike outs (K) and walks (BB). We graph the career trajectory of K rate and BB rate for Clemens (Figure 2) and note his career average values
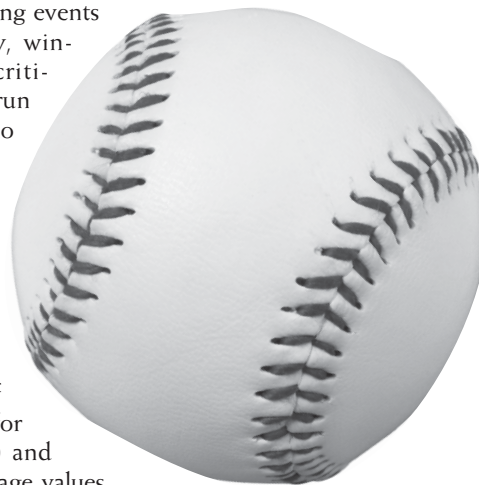
of these statistics are roughly .23 and .078 for K rate and BB rate, respectively.

Again, we see Clemens' strong start, a gradual decline in BB rate as he entered the "first twilight" of his career, followed by a marked improvement. His strikeout rate is more erratic, but, roughly, he improved in his early career, then declined, and then rose again, peaking at the age of 35 in 1998, in his second year with Toronto.

To put these career trajectories into an appropriate context, we require a comparison group. Our first effort was a handful of star-level contemporaries, including Greg Maddux, Johnson, and Schilling. We also include here Ryan, as he was compared to Clemens in the Hendricks report. Their career trajectories for K rate and BB rate are graphed in Figures 3–6.

The career trajectories for Clemens' star contemporaries are nicely fit with quadratic curves. In terms of performance, the curves clearly show steady improvement as the players entered their primes, followed by a marked decline in their strikeout rate (except Ryan, whose K rate trajectory is fairly steady) and a leveling off in their walk rates. The contrast with Clemens' career trajectory is stark. The second act for Clemens is unusual when compared to these other greats because his later success follows an unprecedented period of relative decline. This leaves us with the following question: How unusual is it for a durable pitcher to have suffered a mid-career decline only to recover in his mid- and late 30s?
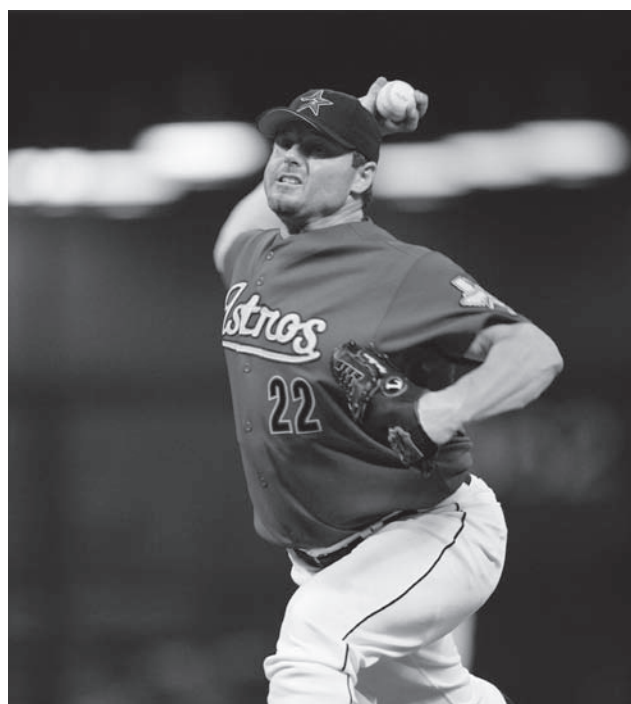
Houston Astros pitcher Roger Clemens throws a pitch against the St. Louis Cardinals during the fifth inning of their Major League game September 24, 2006, in Houston.
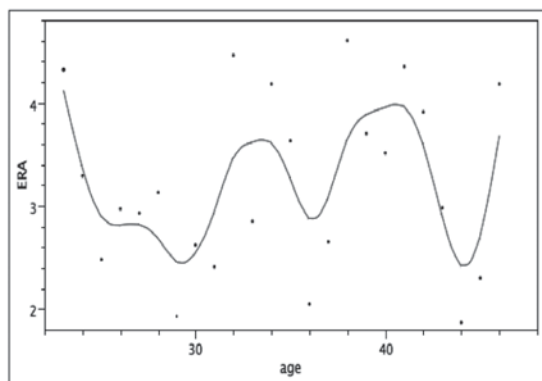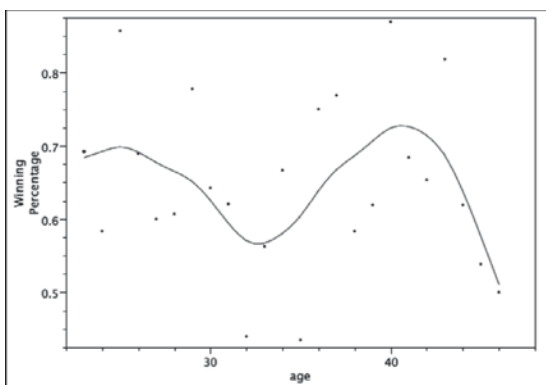(AP Photo/David J. Phillip)

Figure 1. Clemens' winning percentage and ERA throughout time
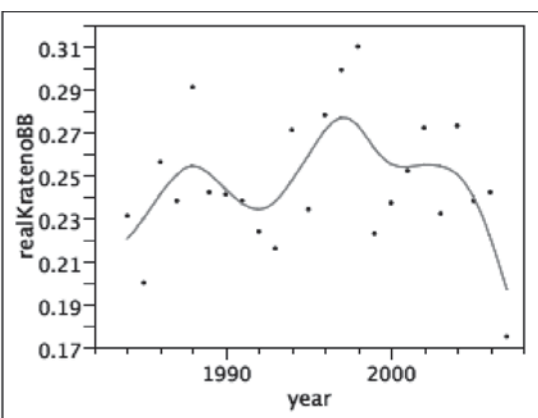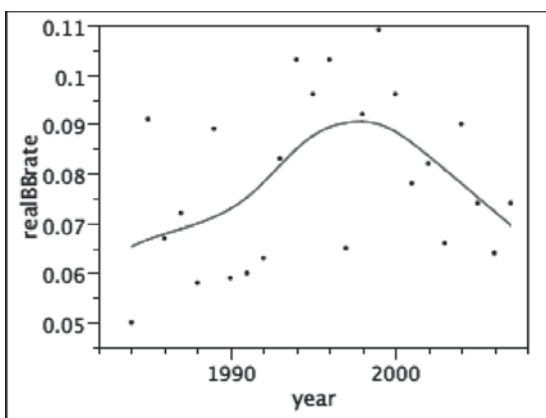
Figure 2. Clemens' BB rate and K rate throughout time
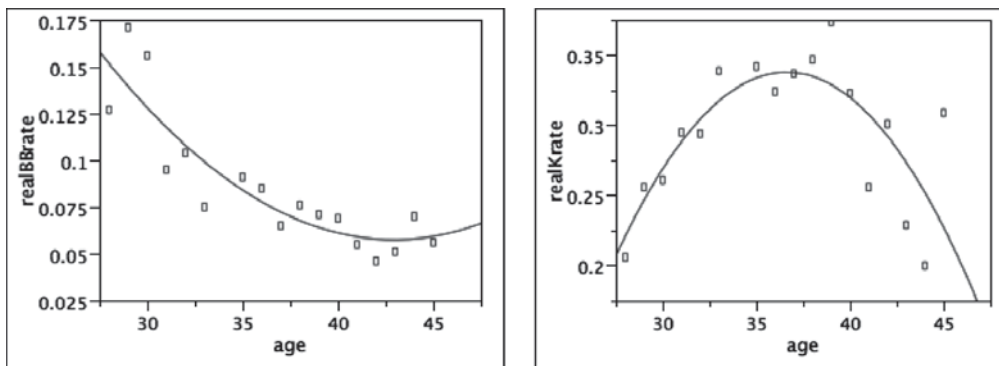
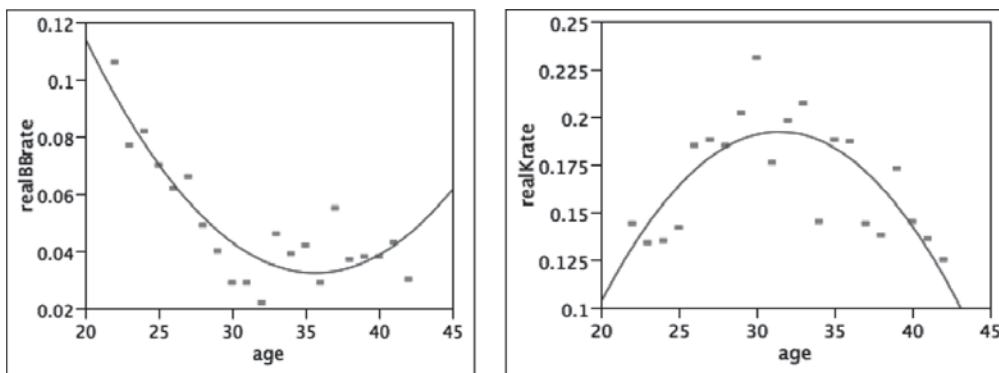Figure 3. Randy Johnson BB rate and K rate throughout time

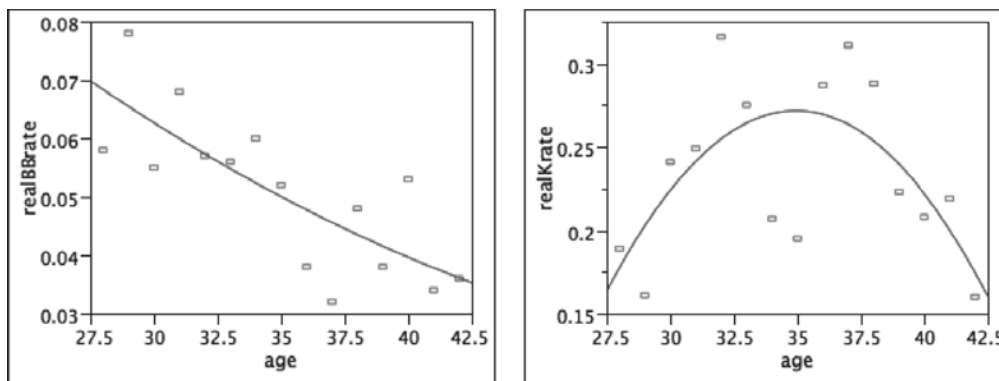Figure 4. Greg Maddox BB rate and K rate throughout time

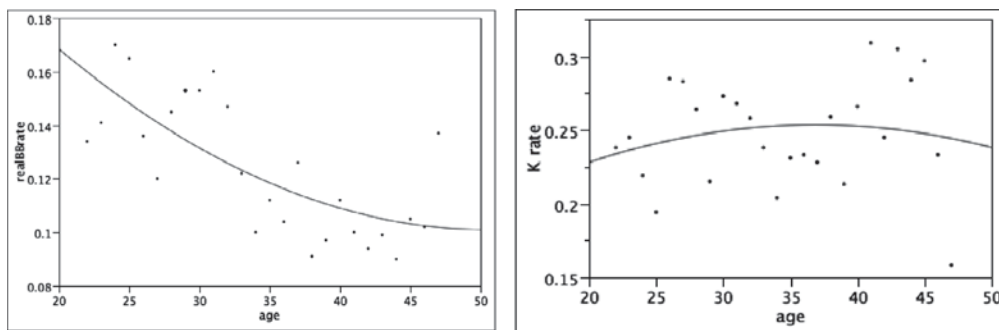Figure 5. Curt Schilling BB rate and K rate throughout time

Figure 6. Nolan Ryan BB rate and K rate throughout time

Database Construction

To perform our statistical analyses, we first obtained data from the Lahman Database, Version 5.5 (*www.baseball1.com*), on all Major League Baseball (MLB) pitchers whose careers were contained in the years 1969–2007. (The starting year of 1969 was selected because of the change in the height of the pitchers' mound, which launched the 'modern era' in baseball.) From that set of pitchers, we constructed a comparison set of all durable starting pitchers by looking at all pitchers who played at least 15 full seasons as a starter (with 10 or more games started per year) and had at least 3,000 innings pitched in those seasons (we note that sensitivity analyses run that included minor perturbations in these criterion indicated the results are quite stable). There were 31 pitchers other than Clemens who fit these criteria. All of these starting pitchers, therefore, had comparably long careers (in years) and innings pitched similar to Clemens. Hence, they were a relevant comparison set, although others could certainly be chosen. See *www.amstat. org/publications/chance* (appendix1) for the names and a set of descriptive statistics for the 31 players and Clemens.

For each pitcher, we looked at the following well-established pitching statistics for each of the years in which they pitched:

WHIP = Walks + hits per inning pitched

BAA = Batting average for hitters when facing the given pitcher

ERA = Earned run average per nine innings pitched

BB Rate = Walk rate

K Rate = Batter strike-out rate per plate appearance (not including walks)

Together, these statistics provide a fairly complete picture of the career trajectory for a starting pitcher.

Trajectory Analyses

To understand and summarize the trajectory each of the five ($j = 1,...,5$) aforementioned statistics take, we fit a quadratic function to each of the 32 ($i = 1,...,32$) focal pitcher's (including Clemens) data at year t, as follows:

$$S_{ijt} = \beta_{0ij} + \beta_{1ij}Age_{it} + \beta_{2ij}Age^2_{ij} + \varepsilon_{ijt} \qquad [1]$$

where S_{ijt} = value of statistic j for pitcher i in their t-th season; Age_{it} = age of pitcher i in their t-th major league season, β_{0ij}, β_{1ij}, and β_{2ij} are an intercept and coefficients describing how Age and Age^2 influence the prediction of the statistics; and ε_{ijt} is a randomly distributed normal error term. As none of the measures studied was near the boundary of their respective ranges, taking transformations (that is standard) had no substantive impact.

We also acknowledge that a quadratic curve may not be the best model for every pitcher's career, including Clemens'. However, the quadratic curve is a simple model with interpretable coefficients that provide a common basis of comparison for all pitchers in our study. The quadratic curve is an appropriate model for the usual trajectory of performance, which expects improvement as a pitcher hits his prime and then decline as he ages (graphically, his performance climbs over and down the proverbial 'hill').

Our goal is not to model the specific trajectory for every player, but to detect those patterns that stick out as highly unusual with respect to a quadratic reference. So, it would
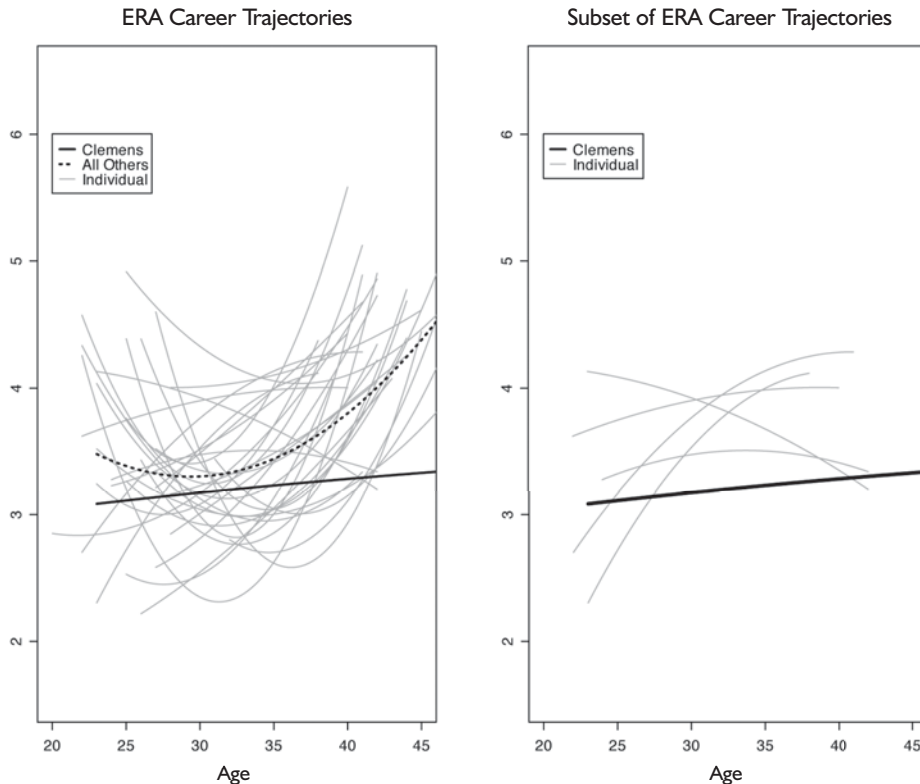
Figures 7a–b

not be appropriate, for example, to consider a cubic fit to each player for our purpose. It is possible that we may not be able to identify interesting patterns for an individual player's trajectory by using only quadratic curves. This is not a concern, however, as we are only interested in determining how often the typical quadratic trajectory occurs among the pool of comparison players.

Our primary interest centers on the coefficient β_{2ij} that describes whether the pitcher's trajectory for that statistic is purely linear ($\beta_{2ij} = 0$), "hump-shaped" ($\beta_{2ij} <0$), or "U-shaped" ($\beta_{2ij} >0$) as he ages. To provide context, one might predict the following patterns, corresponding *a priori* to a pitcher hitting a mid-career 'prime' and then falling off near the end of his career:

WHIP ($\beta_{2ij} >0$)
BAA ($\beta_{2ij} >0$)
ERA ($\beta_{2ij} >0$)
BB Rate ($\beta_{2ij} >0$)
K Rate ($\beta_{2ij} <0$)

Note the sign change for K rate for β_{2ij} as more strikeouts is better, while a lower value for the other statistics is better. Figures 7a and 7b contain a more detailed analysis of the data from the Hendricks report, using ERA. We first present in Figure 7a the ERA curves for the 32 relevant players (31 pitchers + Clemens). Each pitcher's trajectory is depicted with a thin, gray curve, except for Clemens', which is depicted with a thick, black curve. Also given is a dotted curve, which is the quadratic trajectory fit to the data for all players except Clemens. Figure 7b contains the players with curves that have
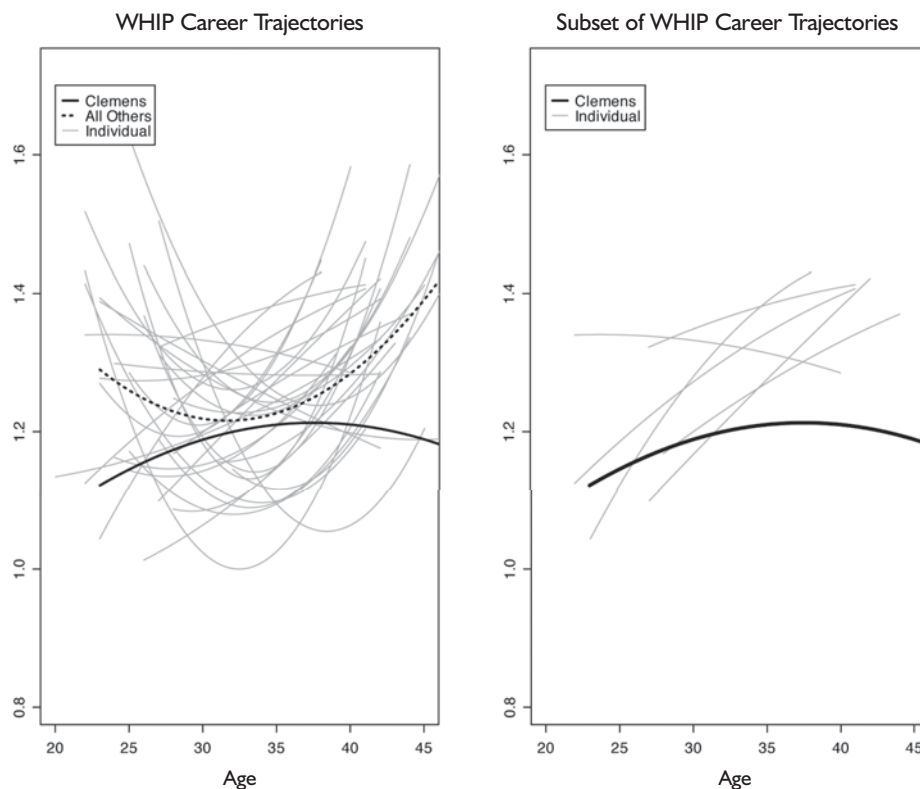
quadratic terms that are 'atypical' ($\beta_{2ij} \leq 0$) compared to the prior hypothesis of a mid-career prime. Six players, including Clemens, have these atypical curves, and, in fact, Clemens' curve looks atypical even within this subset of six players.

Figures 8a and 8b contain career trajectories of WHIP for the same 32 players. Clemens is again within a small subset of seven pitchers who show atypical career paths. Further inspection of his WHIP curve suggests he was the only pitcher to get worse as his career went on and then improve at the end of his career.
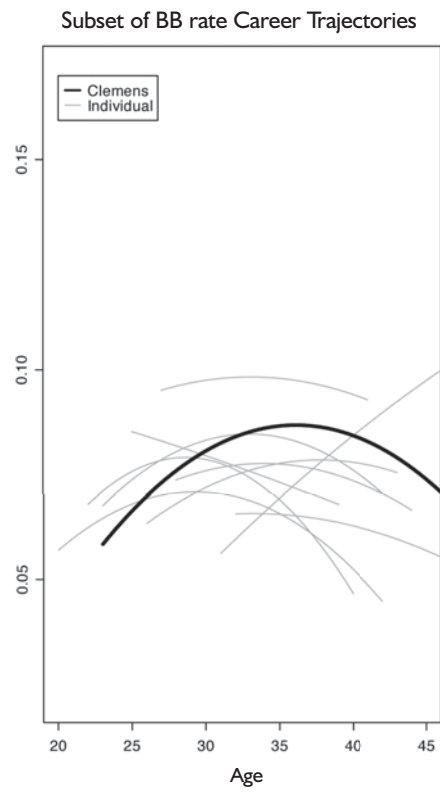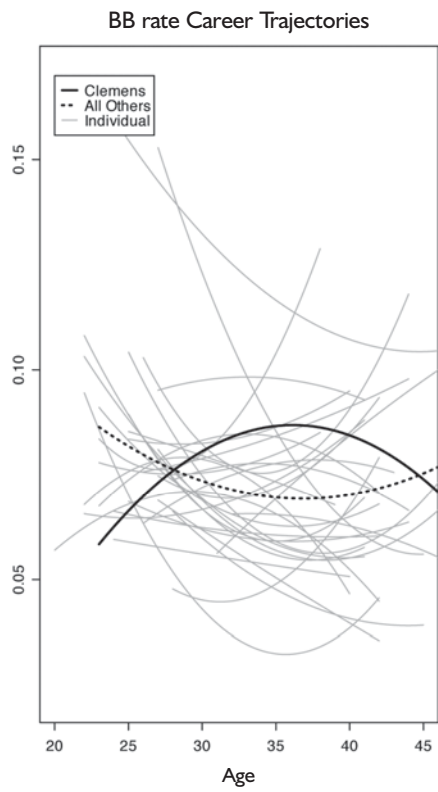
Two additional analyses we performed using ERA and WHIP were to compute the same figures as Figures 6 and 7, but instead using ERA margin and WHIP margin, defined as the difference between the individual ERA and the league average. In the graphs at *www.amstat.org/publicatioins/chance* (Appendix2), we show the ERA margin and WHIP margin curves for Clemens and for the average over the 31 other pitchers. We see little difference between the raw curves in Figures 7 and 8 and the margin curves.

Figures 9a and 9b contain career trajectories of BB rate (walks per batter faced) for the same 32 players. For BB rate, we note there are 10 pitchers who have "inverted-U" fits to their data, with Clemens being one of them. Furthermore, the 'steepness' of his improvement is particularly noticeable in the later years, even among this set of 10.
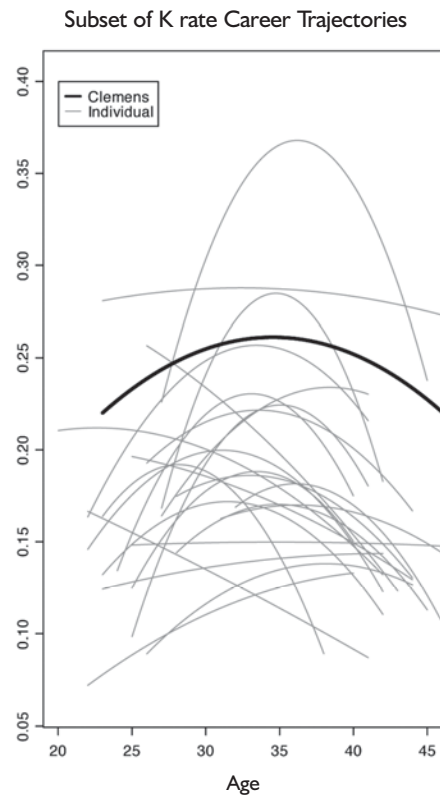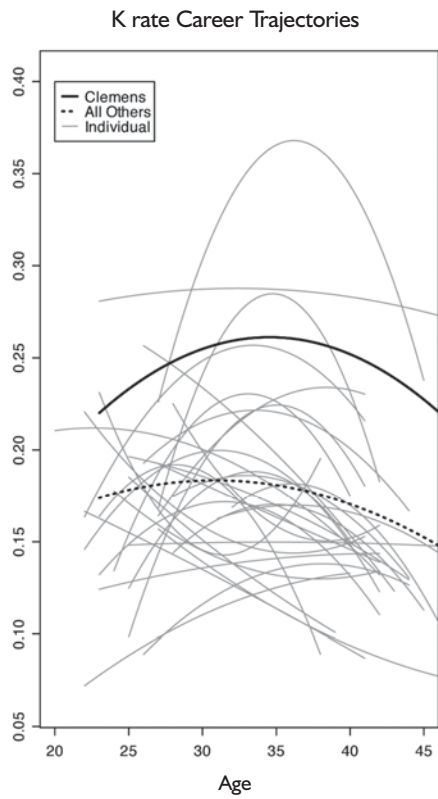
There are several pitching measures for which Clemens' career trajectory does not look atypical, which is the central assertion of the Hendricks report. In Figures 10a and 10b, we give the strikeout rate (K per non-BB batters faced) for each of the 32 durable starting pitchers. Clemens does have an overall higher K rate than most pitchers in this set, but his career
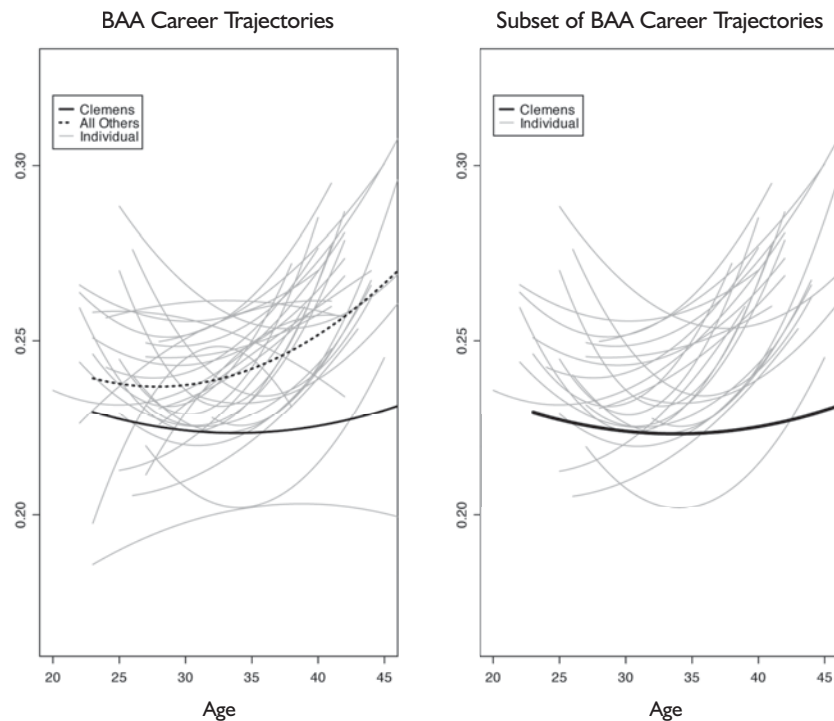

Figures 8a–b

BB rate Career Trajectories

Subset of BB rate Career Trajectories

Figures 9a–b

K rate Career Trajectories

Subset of K rate Career Trajectories

Figures 10a–b

BAA Career Trajectories   Subset of BAA Career Trajectories

Figures 11a–b

trajectory follows a similar shape ($\beta_{2ij} < 0$) to 24 of the other 31 players, at least with respect to the quadratic fit.

In Figures 11a and 11b, we examine BAA (batting average against) for each of the 32 pitchers. Similar to K rate, we again observe that Clemens has a typical shape to his career trajectory to most (24 out of 31) of the other starting pitchers, albeit his curve is somewhat flatter.

Through the use of simple exploratory curve fitting applied to a number of pitching statistics, and for a well-defined set of long-career pitchers, we assessed whether Clemens' pitching trajectories were atypical. Our evidence is suggestive that while most long-term pitchers have peaked mid-career and decline thereafter, Clemens (for some key statistics) worsened mid-career and improved thereafter.

There are many other ways to approach this question, and we expect other researchers will try different techniques. We warn these brave souls that baseball statistics are extraordinarily variable. For example, it is generally assumed that the league average ERA for the National League is lower than that for the American League. This is true—the average gap is 0.25 runs (in favor of the National League). But, in some years, that gap is huge (0.75 runs in 1996), and, in other years, the gap is negligible (nearly 0 in 2001 and 2007). So, while it is tempting to 'control' for patterns such as these, you may just be adding noise to your data by subtracting a random quantity (league-wide statistics) from another.

So, what can we conclude? We can conclude that Clemens' pitching career was atypical for long-term pitchers in terms of WHIP, BB rate, and ERA. In particular, Clemens shows a mid-career decline followed by an end-of-career improvement that is rarely seen. This is a trajectory not seen at all among the comparison group of pitchers identified by Hendricks Sports Management. We emphasize that our analysis is entirely exploratory—we do not believe there exists a reasonable

probability model to test relevant hypotheses by calculating significance levels. The data does not exonerate (nor does it indict) Clemens, as an exploratory statistical analysis of this type never proves innocence or guilt. After analyzing this data set, there are at least as many questions remaining as before. **C**

Further Reading

http://en.wikipedia.org/wiki/Baseball_statistics#Pitching_statistics

Albert, J. (2006) "Pitching Statistics, Talent and Luck, and the Best Strikeout Seasons of All-Time." *Journal of Quantitative Analysis in Sports*, **2**(1):**Article 2**.

Bradlow, E.; Jensen, S.T.; Wolfers, J.; and Wyner, A. (2008) "Report Backing Clemens Chooses Its Facts Carefully." *The New York Times*. February 10, 2008.

Fainaru-Wada, M. and Williams, L. (2007) *Game of Shadows: Barry Bonds, BALCO, and the Steroids Scandal That Rocked Professional Sports*. Gotham Books: New York, New York.

Hendricks, R.A.; Mann, S.; and Larson-Hendricks, B.R. (2008) "An Analysis of the Career of Roger Clemens." *www.rogerclemensreport.com*, January 18.

Mitchell, G.J. (2007) "Report to the Commissioner of Baseball of an Independent Investigation Into the Illegal Use of Steroids and Other Performance Enhancing Substances by Players in Major League Baseball." DLA Piper US LLP: December 13. (Accessed at *http://mlb.mlb.com/mlb/news/mitchell/index.jsp*.)

Silverman, M. (1996) "Baseball End of an Era: No Return Fire from Sox Brass Tried to Keep Ace." *Boston Herald*, December 14, Page 40.

Vecsey, G. (1968) "Baseball Rules Committee Makes 3 Decisions to Produce More Hits and Runs." T*he New York Times*, December 4, Page 57.