Appalachian State University

Department of Economics Working Paper

Number 13-15 | August 2013

Evaluation of Talent in a Changing World: The Case of Major League Baseball

Peter A. Groothuis
Appalachian State University

Kurt W. Rotthoff
Seton Hall University

Mark C. Strazicich
Appalachian State University

Department of Economics
Appalachian State University
Boone, NC 28608
Phone: (828) 262-2148
Fax: (828) 262-6105
www.business.appstate.edu/economics

"Evaluation of Talent in a Changing World: The Case of Major League Baseball"

Peter A. Groothuis
Walker College of Business
Appalachian State University
Boone, NC 28608

Kurt W. Rotthoff
Seton Hall University
Stillman School of Business
South Orange, NJ 07079

Mark C. Strazicich
Department of Economics
Appalachian State University
Boone, NC 28608

July 31, 2013

Abstract:
Comparing talent across time is difficult as productivity changes. To compare talent across time we utilize Major League baseball data from 1871-2010 and time series techniques to determine if the mean and standard deviation of five performance measures are stationary and if structural breaks exist. We identify two structural breaks in the mean slugging percentage: in 1921, the free swinging era of Babe Ruth, and in 1992, the steroid era. Given that productivity changes over time, we develop a simple benchmark technique to compare talent over time and identify superstars. Applications of this measure outside of baseball are also suggested.

JEL Classifications: J24, C22

Keywords: Benchmarking, Major League Baseball, Technology Changes, Structural Breaks

1. Introduction

Games change. New innovations are developed. In many sports it is the equipment that drives the change in the game such as innovation in tennis rackets or in golf club technologies. In other sports it might be the development of a new defensive technique, a new way to swing the bat, throw a pitch, shoot a basket, or hold a putter. When players develop successful innovations they are mimicked and the game changes.

When a game changes comparing talent across different time periods becomes increasingly difficult. To overcome this difficulty we propose a simple benchmarking technique to address the question: How good are players when benchmarked to their peers? Our technique is common in finance, where performance of an asset is not simply measured by the absolute return, but the return relative to some benchmark. In such cases, the benchmark is established as a market portfolio or Security Market Line (Roll 1978) where the portfolio manager's goal is to 'beat the market'. Similar benchmarking is used in many other ways. For example, salaries are benchmarked to relative pay; technological development, research output, and teaching performance, among others, can be similarly benchmarked. We expand the use of such benchmarking to analyze performance when innovations occur.

The use of benchmarking allows for more accurate analysis of performance. For example, benchmarking can be expanded to provide more detailed analysis of talent levels; not just of today's performances, but for a viable measure of talent across different eras where different innovations have occurred.[1] Talent is highly valued, thus accurate measures of relative talent today, and comparisons across time, are also highly valued.

---

[1] As an example: a firm comparing productivity of an employee pre- and post-computer era would be an inaccurate measure of relative productivity.

In some industries, however, talent is difficult to measure. This measurement difficulty increases as you measure the talent within an industry over long periods of time, for instance from 1871-2010. This problem is also complicated by the fact that when the opportunity to reveal talent is limited, true talent does not have the opportunity to reveal itself (Terviö 2009).

Given that talent changes over time, having a moving benchmark reveals more accurate measurement of the true talent levels of a player at any given point of time. Throughout the years technologies, skills, strength, and training have changed. Having a benchmark increases the accuracy of measurement and the ability to truly compare talent over time. Moreover, having a benchmark provides a convenient method to identify superstars. In addition, having a benchmark may provide insight to help identify important innovative players who changed the game. To determine if a benchmark measure of performance changes we use time series techniques on the mean and standard deviation of several traditional performance measures in baseball. We find that most of these time series are stationary around one or two structural breaks. Perhaps most noteworthy, we find two structural breaks in the slugging percentage mean, in 1921 and 1992, that correspond with the early years of the Babe Ruth free swinging era and the steroid era.

In Section two, we discuss the data and the tests used to identify the benchmark and when the game changes. In Section three, we evaluate talent both by comparing players to an absolute standard and by comparing players to a changing benchmark. We conclude in Section four.

2. Data and Structural Breaks

Major League Baseball attracts the best baseball players in the world. The first professional baseball team was established in 1869 (the "Cincinnati Red Stockings"). The league started in the late 1800s and continues today. Superstars have commonly been identified by the record books. However, we identify superstars in baseball by their deviation from the mean. Using data from Sean Lahman's Baseball Database on all players from 1871-2010 with at least 100 at-bats, we measure slugging percentage (SLUG), home runs per hundred at bats (HR), batting average (BAVE), and runs batted in per hundred at bats (RBI).[2] With 35,728 single season observations we find that the average player hit 7 homeruns per season (with a maximum of 73), had 42.5 runs batted in (RBI), and a slugging percentage of .379. To create a time series to identify structural breaks we calculate both the mean and standard deviation of each performance measure for each season. This provides annual time series from 1871-2010 that consist of 140 seasonal observations for each series.

To determine if the time series measures of player performance are stationary (i.e., a deterministic trend) or non-stationary (i.e., a stochastic trend) and to identify structural breaks, we utilize the one- and two-break minimum LM unit root tests proposed by Lee and Strazicich (2003, 2004). Following Perron (1989), it is well known that ignoring an existing structural break in unit root tests will reduce the ability to reject a false unit root null hypothesis.[3] To overcome this drawback, Perron proposed including dummy variables in the usual augmented Dickey-Fuller unit root test (ADF test) to allow

[2] Sean Lahman's Baseball Database: http://baseball1.com/2011/01/baseball-database-updated-2010/. Slugging percentage is calculated as total bases divided by the number of at-bats.

[3] By "structural break," we imply a significant but infrequent, permanent change in the level and/or trend of a time series. See Enders (2010) for additional background discussion on structural breaks and unit root tests.

for one known, or "exogenous," structural break. In subsequent work, Zivot and Andrews (1992, ZA hereafter), among others, proposed unit root tests that allow for one unknown break to be determined "endogenously" from the data. The ZA test selects the break where the *t*-statistic testing the null of a unit root is minimized (i.e., the most negative). The ZA test, however, and other similar ADF-type endogenous break unit root tests derive their critical values assuming no break under the null hypothesis. Nunes, Newbold, and Kuan (1997) and Lee and Strazicich (2001), among others, show that this assumption can lead to spurious rejections of the unit root hypothesis in the presence of a unit root with break. As a result, when using these tests, researchers can incorrectly conclude that a time series is "trend-break stationary" when in fact the series has a unit root with break. To avoid these drawbacks, we utilize the one- and two-break minimum LM unit root tests developed by Lee and Strazicich (2003, 2004). The endogenous LM unit root test has the desirable property that its test statistic is not subject to spurious rejections. Thus, conclusions are more reliable since rejection of the null hypothesis unambiguously implies that the series is stationary around one or two breaks in the level and/or trend.

Our testing methodology can be summarized as follows.[4] According to the LM "score" principle, the test statistic for a unit root can be obtained from the following regression:

$$\Delta y_t = \delta'\Delta Z_t + \phi\widetilde{S}_{t-1} + \Sigma\gamma\Delta\widetilde{S}_{t-i} + \varepsilon_t, \qquad (1)$$

where $\widetilde{S}_t = y_t - \widetilde{\psi}_x - Z_t\widetilde{\delta}$, $t=2,..,T$; $\widetilde{\delta}$ are the coefficients from the regression of Δy_t on ΔZ_t and $\widetilde{\psi}_x$ is the restricted MLE of ψ_x $(\equiv\psi + X_0)$ given by $y_1 - Z_1\widetilde{\delta}$. $\Delta\widetilde{S}_{t-i}$ terms are included

[4] Gauss codes for the one- and two-break minimum LM unit root test are available on the web site http://www.cba.ua.edu/~jlee/gauss.

as necessary to correct for serial correlation. ε_t is the contemporaneous error term and is assumed to be independent and identically distributed with zero mean and finite variance. Z_t is a vector of exogenous variables contained in the data generating process. Z_t is described by $[1, t, D_{1t}, D_{2t}, DT_{1t}^*, DT_{2t}^*]'$, where $D_{jt} = 1$ if $t \geq T_{Bj} + 1$, $j = 1,2$, and zero otherwise, $DT_{jt}^* = t$ if $t \geq T_{Bj} + 1$, and zero otherwise, and T_{Bj} is the time period of the structural break. Note that the testing regression (1) involves ΔZ_t instead of Z_t so that ΔZ_t is described by $[1, B_{1t}, B_{2t}, D_{1t}, D_{2t}]'$, where $B_{jt} = \Delta D_{jt}$ and $D_{jt} = \Delta DT_{jt}^*$, $j=1, 2$. Thus, B_{1t} and B_{2t}, and D_{1t} and D_{2t}, correspond to structural changes or breaks in the level and trend under the alternative, and to one period jumps and permanent shifts in the level under the null hypothesis, respectively. The unit root null hypothesis is described by $\phi = 0$ and the LM test statistic is defined by:

$$\tilde{\tau} \equiv t\text{-statistic testing the null hypothesis } \phi = 0. \qquad (2)$$

To endogenously determine the location of two breaks ($\lambda_j = T_{Bj}/T$, j=1, 2), the minimum LM unit root test uses a grid search to determine the combination of two break points where the unit root test statistic is minimized. Since the critical values for the model with trend-break vary (somewhat) depending on the location of the breaks (λ_j), we employ critical values corresponding to the identified break points.

To determine the number of lagged augmented terms $\Delta\tilde{S}_{t-i}$, $i = 1,..,k$, that are included to correct for serial correlation, we employ the following sequential "general to specific" procedure. At each combination of two break points $\lambda = (\lambda_1, \lambda_2)'$ over the time interval [.1T, .9T] (to eliminate end points) we determine k as follows. We begin with a maximum number of k = 8 lagged terms and examine the last term to see if its t-statistic is significantly different from zero at the 10% level (critical value of 1.645 in an asymptotic normal distribution). If insignificant, the $k = 8$ term is dropped and the model is re-estimated using $k = 7$ terms, etc., until the maximum lagged term is found, or $k = 0$. Once the maximum number of lagged terms is found, all lower lags remain in the regression. The process is repeated for each combination of two break points to jointly identify the breaks and the test statistic at the point where the unit root test statistic is minimized.[5]

The LM unit root test results are reported in Table 1. In each case, we begin by applying the two-break LM unit root test. If only one break is identified (at the 10% level of significance) in the two-break test, we re-examine the series using the one-break LM unit root test. The mean slugging percentage (SLUGM) rejects a unit root at the 5% significance level, implying that SLUGM is a stationary series with two level and trend breaks in 1921 and 1992. For the slugging percentage standard deviation (SLUGSD), only one structural break was significant in the two-break test. We therefore re-tested this series using the one-break test. In contrast to SLUGM, the SLUGSD cannot reject a unit root at the 10% level of significance, implying that this series is nonstationary. The unit root hypothesis cannot be rejected for the homerun mean (HRM) at the 10% level of significant, implying that this series is nonstationary and has a positive stochastic trend. In contrast, the unit root hypothesis is rejected for the homerun standard deviation (HRSD) at the 5% level of significance, implying that this series is stationary with two level and trend breaks in 1920 and 1966. The mean batting average (BAVEM) cannot reject the unit root hypothesis at the 10% level of significance, implying that BAVEM is

[5] This type of method has been shown to perform better than other data-dependent procedures to select the optimal k (e.g., Ng and Perron, 1995).

a nonstationary series. In contrast, the batting average standard deviation (BAVESD) rejects the unit root hypothesis at the 1% level of significance, implying that BAVESD is a stationary series with two breaks in 1906 and 1933. The mean of runs batted in (RBIM) rejects the unit root hypothesis at the 10% level of significance, implying that RBIM is a stationary series with one level and trend break in 1887. The standard deviation of runs batted in (RBISD) rejects the unit root hypothesis at the 5% level of significance, implying that RBISD is a stationary series with one level and trend break in 1921.

We next perform regressions on the identified level and trend breaks for the five performance time series that reject the unit root hypothesis (SLUGM, HRSD, BAVESD, RBIM, and RBISD). Note that regressions will not be undertaken for SLUGSD, HRM, and BAVEM, since the results in Table 1 indicate that these series are nonstationary and spurious regressions can occur. In reporting our results, the coefficients on the first D_t and T_t terms denote the intercept and trend slope in the time period from 1871 to the first break, the time period after the first break to the second break, and the time period after the second break to 2010, respectively. In each regression we correct for serial correlation by including lagged values of the dependent variable as necessary using the “general to specific” approach described for the LM unit root tests. White’s robust standard errors are used to correct for heteroskedasticity.

We begin by regressing the slugging percentage mean (SLUGM) on the two level and trend breaks identified in Table 1. The results are reported in Table 2. Following each break, there is a significant increase (upward shift) in the mean slugging percentage. Perhaps most interesting is that the1921 break coincides with the early years of the Babe Ruth era and the 1992 break coincides with the steroid era often associated with Mark

McGwire, Sammy Sosa, and Barry Bonds, among others. Following each break there is a slight downward trend in SLUGM, which is statistically significant only in the time period of 1992-2010.

We next examine the regression on the two level and trend breaks identified in Table 1 for the standard deviation of home runs (HRSD). The results are reported in Table 3. Following each break, there is a significant upward shift in the standard deviation of home runs indicating that the dispersion in home run performance increased after each break. Again, we see that the first break (in 1920) is associated with the early years of the Babe Ruth era. In each time period there is a positive and significant trend, which steepens after 1920. Then, following the second break (in 1966), the trend slope remains positive but flattens somewhat. These results suggest that Babe Ruth had a significant impact on the game that lead to a greater dispersion in home run performance among players.

We next examine regression results for the batting average standard deviation (BAVESD), which was found to be stationary around two structural breaks in 1906 and 1933. The results are reported in Table 4. Following each break, there is a significant downward shift in the standard deviation indicating that the variation in batting average decreased after 1906 and 1933. There is a negative trend in each time period, while only the trend in the last period (1934-2010) is statistically significant.

We next examine results for our final performance measures of the mean and standard deviation of runs batted in (RBIM and RBISD) in Table 5 and 6. Following the break in 1887 the RBIM increases. While there is a slight negative trend after the break, the trend slope is not significant (p-value > 0.1). For RBISD, following the break in 1921 there is a slight increase in RBISD. There is a negative trend in each time period, but the trend slope is not significant in either period. While the break in RBISD is again associated with the Babe Ruth era, there is little change in the series. Overall, the results for RBI suggest that there has been little change in the RBIM and RBISD throughout 1871-2010.

To better visualize the regression results reported in Tables 2-6, we construct simple plots of the estimated trends and the actual data in Figures 1-5. As in the regression results, perhaps most interesting is the plot of the slugging percentage mean (SLUGM) displayed in Figure 1. From Figure 1, we can easily observe the significant upward shifts in SLUGM that occurred in 1921 and 1992, with the biggest increase apparent in 1921 during the early years of the Babe Ruth era. Similarly, in Figure 2, we see a significant upward shift in the standard deviation of home runs (HRSD) in 1920. In Figure 3, we can observe the general decline in the batting average standard deviation (BAVESD). In Figures 4 and 5, we observe the relative overall stability of the mean and standard deviation of runs batted in (RBIM and RBISD), respectively. Given that we did not perform regressions for the slugging percentage standard deviation (SLUGSD), home runs mean (HRM), and batting average mean (BAVEM) due to the inability to reject a unit root for these series, we provide plots of these series only for the actual data. The plots are reported in Figures 6-8. While these plots are provided for convenience, it is difficult to provide analysis relevant to constructing a baseline of baseball performance for these series since they could not reject a unit root and resemble random walks.[6]

[6] In a unit root process, a structural break in the level can be interpreted as an unusually large one-time shock or outlier, while a break in the trend can be interpreted as a permanent change in the drift.

The above results suggest that major league baseball performance had a structural break around 1920-1921. The slugging percentage mean (SLUGM) had a positive break in 1921 suggesting that the average player started to hit for power and hit more doubles, triples and home runs. The runs batted in standard deviation (RBISD) had a positive structural break in 1921 as did the home run standard deviation (HRSD) in 1920. Both of these structural breaks suggest that players after 1920 or 1921 became more diverse in their performance measures with some players hitting for power while others did not. This result might suggest that with the success of Babe Ruth's free swinging style others that could mimicked his innovation and hit for power as well. Most notable among the other breaks is the 1992 positive break in the slugging perecentage mean (SLUGM) that is closely associated with the steriod era. The existence of breaks and (stationary) deterministic trends for most baseball performance series suggest that adopting a benchmark technique is the best way to evaluate talent over time.

3. Benchmarking

The deliberation on superstars and their relative performance is oft debated and hard to measure, particularly when the comparison happens over different periods of time. When structural breaks occur in the game it makes accurate comparisons nearly impossible. A more accurate way to measure these talents across time allows for more accurate identification of truly great stars. Given a seemingly endless set of debates and lists of superstars we propose a measurement technique, adopted from the finance literature, to compare stars in a relative performance measure of their same generational cohort.

In finance, all funds want their returns to be positive; however, the true measure of success is the ability to 'beat the market'. Thus, the overall ranking is relative to some moving measure over time, referred to as the market average. Applying this relative measure to sporting events will allow us to compare groups of individuals that may have played in very different eras.

In Tables 7-10, we report the means of batting averages, slugging percentage, home runs per hundred at bats, and runs batted in per hundred at bats, respectively. In each table we report the top ten talented players as measured in absolute terms by the overall standard deviations above the overall mean, and the benchmark measure as the yearly standard deviation above the yearly mean. The first measure treats the entire population as peers and does not account for changes in the game. The second technique compares talent directly to their peers.

In Table 7 we report the ten players with the best batting average. We find that using the absolute measure the ten best players all occur in the early years of baseball with eight of the ten in the late 1800s, one in 1901, and the last, Roger Hornsby, in 1924. Using the benchmarked measure we find that the ten best players come from all eras in baseball. Manny Ramirez is the most recent hitting 3.75 standard deviations above the season mean. Other notables on this list are Ted Williams in 1941, George Brett in 1980 and Tony Gwynn in 1994.

Next we report the results of the slugging percentage for both measures of talent in Table 8. Using the absolute standard, Babe Ruth makes the top ten lists four times and Barry Bonds makes the list 3 times. The other three making the top ten are Lou Gehrig, Roger Hornsby, and Mark McGwier. On the benchmark list Babe Ruth makes the list 5 times with the top two in 1920 and 1921 where the structural break in the mean of the slugging percentage series occurs. Barry Bond makes the top ten lists in 2001, 2002 and 2004. He makes the list many years after the second structural break in 1992. Other players that make the list are Lou Gehrig in the eighth position and Ted Williams in the ninth.

We next turn our attention to home runs. We report the means in Table 9. First we find that using the absolute standard Barry Bonds and Mark McGwier dominate the list of the top 13 with Bonds being in the first position hitting home runs 7.37 standard deviations above the mean and making the list three times and McGwier making the list six times in the second through seventh position. Babe Ruth only makes the list in the tenth position in 1920 hitting 5.45 standard deviations above the absolute mean.

When you take the home runs per at-bats for each individual player for a given year, and rank the standard deviations above the mean for each given year, the top ranked home run hitter is Babe Ruth in 1920 (Yankees), 1921 (Yankees), 1919 (Boston), and 1927 (Yankees). He was 10.58, 8.07, 7.26, and 7.04, respectively, standard deviations above the mean. The fifth highest ranked player is Ned Williamson (1884 Chicago), followed by Ruth (1926), Ruth (1924), Buck Freeman (1899 Washington Senators), Ruth (1928), and Gavvy Cravath (1915 Phillies). From the modern era the highest ranked players are Barry Bonds (2001 San Francisco), in 13th place, at 5.85 standard deviations above the mean and Mark McGwier (1998 and 1997 St Louis), in 19th and 20th place, at 5.4 standard deviations above the mean.

Babe Ruth, in his 1920 playing year with the New York Yankees, was 10.58 standard deviations above the mean for that season. This is simply amazing and displays his level of performance relative to the competition he faced. To put this in perspective, if Babe Ruth was 10.58 standard deviations above the mean in 2001, when Barry Bonds set the single season home run record, and had the same 476 at-bats that Barry Bonds did, he would have hit 120 home runs. Barry Bond still holds the single season mark with 73.

Next, we measure the RBIs per at-bat of the players throughout time to measure how each player performs relative to the mean of the year they played in, again with at least 100 at-bats.

In Table 10, we report the results of the superstars as measured by standard deviations above the mean. We find that Reb Russell, playing for the Pittsburgh Pirates, has the highest ranking in RBIs both using the absolute and the benchmark standards.[7] He was 5.04 above the absolute mean and 4.93 standard deviations above the season mean. Other notable players on the absolute standard list are Babe Ruth in 1921 in the sixth position, Manny Ramirez in 1999 in the seventh position and Mark McGwier in 2000 in the ninth position. Using the benchmark standard Babe Ruth has 5 of the top ten rankings. Babe Ruth ranks third, fourth, fifth, sixth and tenth. No players from the modern era make the top ten. We do find that Manny Ramirez is ranked 13th and 20th as the highest ranked modern era player.

4. Conclusion

When innovation occurs players/workers who mimic the innovation also receive an increase in productivity. As this productivity changes, it becomes increasingly difficult for management to compare productivity over different periods of time. When this occurs, relative measures have more value. We propose that the use of a benchmark

---

[7] Reb Russell was a pitcher from 1912-1917 with the Chicago White Soxs. He did not become a big hitter until after developing arm troubles and finding his hitting in the minor leagues.

measure is more accurate in finding superstar players in baseball and this strategy can be used by managers when analyzing superstar employees over different periods of time.

Using various hitting performance measures our time series analysis identifies two major structural breaks in performance measures, in about 1920 and 1992. We suggest that the game of baseball significantly changed during these times. Using benchmarking techniques we find that Babe Ruth was the best power hitter in baseball compared to his peers particularly just prior to the structural break in the game. We suggest that the structural break occurs when players began to mimic Babe Ruth's technique. After focusing on both the structural breaks and the benchmark measure of talent, we suggest that Babe Ruth was not only the best power hitter compared to his peers but he also changed the game. The other structural change came at, what could be argued to be, the beginning of the steroid era.

References

Enders, Walter (2010). *Applied Econometric Time Series*, 3rd edition. Hoboken, NJ: John Wiley & Sons, Inc.

Lee, J., and M. C. Strazicich (2001). "Break Point Estimation and Spurious Rejections with Endogenous Unit Root Tests," *Oxford Bulletin of Economics and Statistics* 63(5): 535-558.

Lee, J., and M. C. Strazicich (2003). "Minimum Lagrange Multiplier Unit Root Test with Two Structural Breaks," *Review of Economics and Statistics* 8(4): 1082-1089.

Lee, J., and M. C. Strazicich (2004). "Minimum LM Unit Root Test," Manuscript, Department of Economics, Appalachian State University.

Ng, S., and P. Perron (1995). "Unit Root Tests in ARMA Models with Data-Dependent Methods for the Selection of the Truncation Lag," *Journal of the American Statistical Association*, 90(429): 269-281.

Nunes, L., P. Newbold, and C. Kuan (1997). "Testing for Unit Roots with Breaks: Evidence on the Great Crash and the Unit Root Hypothesis Reconsidered," *Oxford Bulletin of Economics and Statistics* 59: 435-448.

Perron, P. (1989). "The Great Crash, the Oil Price Shock, and the Unit Root Hypothesis," *Econometrica* 57: 1361-1401.

Roll, R. (1978). "Ambiguity when Performance is Measured by the Securities Market Line." *Journal of Finance* 33(4), 1051-69.

Terviö, Marko (2009). "Superstars and Mediocrities: Market Failure in The Discovery of Talent." *Review of Economic Studies* 72(2), 829-850.

Zivot, Eric and Donald W. K. Andrews (1992). "Further Evidence on the Great Crash, the Oil-Price Shock and the Unit Root Hypothesis," *Journal of Business and Economic Statistics*, 10: 251-270.

Table 1. LM Unit Root Test Results, 1871-2010

Time Series	*k*	*Breaks*	*Test Statistic*	*Break Points*
SLUGM	0	1921, 1992	-5.714**	λ = (.4, .8)
SLUGSD	1	1904	-3.806	λ = (.2)
HRM	0	1949, 1975	-5.094	λ = (.6, .8)
HRSD	0	1920, 1966	-6.156**	λ = (.4, .6)
BAVEM	6	1891, 1941	-5.151	λ = (.2, .6)
BAVESD	0	1906, 1933	-7.344***	λ = (.2, .4)
RBIM	8	1887	-4.397*	λ = (.2)
RBISD	5	1921	-4.707**	λ = (.4)

Notes: SLUG, HR, BAVE, RBI, and ERA denote annual slugging percentage, homeruns, batting average, runs batted in, and earned run average of all players in the series, where M denotes the mean and SD denotes the standard deviation, respectively. The Test Statistic tests the null hypothesis of a unit root, where rejection of the null implies a trend-break stationary series. *k* is the number of lagged first-differenced terms included to correct for serial correlation. The critical values for the one- and two-break LM unit root tests come from Lee and Strazicich (2003, 2004). The critical values depend on the location of the breaks, $\lambda = (T_{B1}/T, T_{B2}/T)$, and are symmetric around λ and (1-λ). *, **, and *** denote significant at the 10%, 5%, and 1% levels, respectively.

Table 2. OLS Regression Results of Slugging Percentage Mean (SLUGM) on Level and Trend Breaks in 1921 and 1992, 1871-2010

$$SLUGM_t = 0.126D_{1871\text{-}1921} + 0.155D_{1922\text{-}1992} + 0.173D_{1993\text{-}2010}$$

(4.650)*** (4.978)*** (5.95)***

$$+ 0.0002T_{1871\text{-}1921} - 0.0001T_{1922\text{-}1992} - 0.0008T_{1993\text{-}2010} + \text{lags}(1) + e_t$$

(1.471) (-1.336) (-2.178)**

Adjusted R-squared = 0.791 SER = 0.016

Notes: Dependent variable is the slugging percentage mean in year t. t-statistics are shown in parentheses. D and T represent dummy variables for the three identified intercepts and trends respectively. White's robust standard errors were utilized to control for heteroskedasticity. Lagged values of the batting average standard deviation were included to correct for serial correlation. ***, **, and * denote significant at the 1%, 5%, and 10% levels, respectively.

Table 3. OLS Regression Results of Home Run Standard Deviation (HRSD) on Level and Trend Breaks in 1920 and 1966, 1871-2010

$HRSD_t = 0.002D_{1871\text{-}1920} + 0.006D_{1921\text{-}1992} + 0.008D_{1993\text{-}2010}$

(3.876)*** (6.154)*** (5.433)***

$+ 0.00003T_{1871\text{-}1920} + 0.00006T_{1921\text{-}1992} + 0.00002T_{1993\text{-}2010} + \text{lags}(1) + e_t$

(1.967)* (3.142)*** (2.233)**

Adjusted R-squared = 0.947 SER = 0.001

Notes: Dependent variable is the home run standard deviation in year t. t-statistics are shown in parentheses. D and T represent dummy variables for the three identified intercepts and trends respectively. White's robust standard errors were utilized to control for heteroskedasticity. Lagged values of the batting average standard deviation were included to correct for serial correlation. ***, **, and * denote significant at the 1%, 5%, and 10% levels, respectively.

Table 4. OLS Regression Results of Batting Average Standard Deviation (BAVESD) on Level and Trend Breaks in 1906 and 1933, 1871-2010

$$BAVESD_t = 0.051D_{1871\text{-}1906} + 0.045D_{1907\text{-}1933} + 0.039D_{1934\text{-}2010}$$

(7.740)*** (7.246)*** (7.653)***

$$- 0.00003T_{1871\text{-}1906} - 0.00007T_{1907\text{-}1933} - 0.00005T_{1934\text{-}2010} + lags(4) + e_t$$

(-0.598) (-1.186) (-4.674)***

Adjusted R-squared = 0.875 SER = 0.002

Notes: Dependent variable is the batting average standard deviation in year t. t-statistics are shown in parentheses. D and T represent dummy variables for the three identified intercepts and trends respectively. White's robust standard errors were utilized to control for heteroskedasticity. Lagged values of the batting average standard deviation were included to correct for serial correlation. ***, **, and * denote significant at the 1%, 5%, and 10% levels, respectively.

Table 5. OLS Regression Results of Runs Batted In Mean (RBIM) on Level and Trend Break in 1887, 1871-2010

$RBIM_t = 0.005D_{1871\text{-}1887} + 0.023D_{1888\text{-}2010}$

(0.263) (2.679)***

$+ 0.002T_{1871\text{-}1887} + 0.00002T_{1888\text{-}2010} + lags(5) + e_t$

(0.989) (0.688)

Adjusted R-squared = 0.639 SER = 0.010

Notes: Dependent variable is the runs batted in mean in year t. t-statistics are shown in parentheses. D and T represent dummy variables for the two identified intercepts and trends respectively. White's robust standard errors were utilized to control for heteroskedasticity. Lagged values of the batting average standard deviation were included to correct for serial correlation. ***, **, and * denote significant at the 1%, 5%, and 10% levels, respectively.

Table 6. OLS Regression Results of Runs Batted In Standard Deviation (RBISD) on Level and Trend Break in 1921, 1871-2010

$$RBISD_t = \underset{(2.930)^{***}}{0.025}D_{1871\text{-}1921} + \underset{(3.388)^{***}}{0.026}D_{1922\text{-}2010}$$

$$- \underset{(-1.040)}{0.00008}T_{1871\text{-}1921} - \underset{(-1.635)}{0.00002}T_{1922\text{-}2010} + lags(4) + e_t$$

Adjusted R-squared = 0.396 SER = 0.004

Notes: Dependent variable is the runs batted in standard deviation in year t. t-statistics are shown in parentheses. D and T represent dummy variables for the two identified intercepts and trends respectively. White's robust standard errors were utilized to control for heteroskedasticity. Lagged values of the batting average standard deviation were included to correct for serial correlation. ***, **, and * denote significant at the 1%, 5%, and 10% levels, respectively.

Table 7: Batting Average: Absolute Standard vs. Benchmark

	Player	Year	SD above the season mean	Rank	Player	Year	SD above the season mean
1	Levi Meyerle	1871	5.659514	1	Bob Hazle	1957	3.86
2	Hugh Duffy	1894	4.368691	2	Manny Ramirez	2008	3.75
3	Tip O'Neill	1887	4.258268	3	Ted Williams	1941	3.69
4	Ross Barnes	1872	4.187384	4	George Brett	1980	3.68
5	Cal McVey	1871	4.164273	5	Tip O'Neill	1887	3.65
6	Ross Barnes	1876	4.095538	6	Tony Gwynn	1994	3.59
7	Nap Lajoie	1901	4.043987	7	Oscar Gamble	1979	3.57
8	Ross Barnes	1873	4.019332	8	Tris Speaker	1916	3.54
9	Willie Keeler	1897	3.977446	9	David Dellucci	1999	3.54
10	Roger Hornsby	1924	3.971277	10	Jack Glasscock	1884	3.51

Table 8: Slugging Percentage: Absolute Standard vs. Benchmark

Rank	Player	Year	SD above the absolute mean	Rank	Player	Year	SD above the season mean
1	Barry Bonds	2001	5.65	1	Babe Ruth	1920	5.77
2	Babe Ruth	1920	5.49	2	Babe Ruth	1921	5.21
3	Babe Ruth	1921	5.45	3	Barry Bonds	2001	5.03
4	Barry Bonds	2004	5.06	4	Barry Bonds	2004	4.91
5	Barry Bonds	2002	4.90	5	Barry Bonds	2002	4.79
6	Babe Ruth	1927	4.59	6	Babe Ruth	1927	4.57
7	Lou Gehrig	1927	4.51	7	Babe Ruth	1926	4.50
8	Babe Ruth	1923	4.50	8	Lou Gehrig	1927	4.49
9	Rogers Hornsby	1925	4.40	9	Ted Williams	1941	4.36
10	Mark McGwier	1998	4.36	10	Babe Ruth	1924	4.35

Table 9: Home Runs: Absolute Standard vs. Benchmark

Rank	Player	Year	SD above the absolute mean	Rank	Player	Year	SD above the season mean
1	Barry Bonds	2001	7.37	1	Babe Ruth	1920	10.58
2	Mark McGwier	1997	6.53	2	Babe Ruth	1921	8.07
3	Mark McGwier	1998	6.51	3	Babe Ruth	1919	7.26
4	Mark McGwier	2000	6.40	4	Babe Ruth	1927	7.04
5	Mark McGwier	1999	5.81	5	Ned Williamson	1884	7.01
6	Mark McGwier	1995	5.71	6	Babe Ruth	1926	6.83
7	Mark McGwier	1996	5.71	7	Babe Ruth	1926	6.50
8	Hill Glenallen	2000	5.62	8	Buck Freeman	1899	6.41
9	Barry Bonds	2004	5.58	9	Babe Ruth	1928	6.11
10	Babe Ruth	1920	5.45	10	Gavvy Cravath	1915	6.08
11	Barry Bonds	2003	5.30	13	Barry Bonds	2001	5.85
12	Frank Thomas	2005	5.24	19	Mark McGwier	1998	5.42
13	Barry Bonds	2002	5.23	20	Mark McGwier	1997	5.41

Table 10: RBIs: Absolute Standard vs. Benchmark

Rank	Player	Year	SD above the absolute mean	Rank	Player	Year	SD above the season mean
1	Reb Russell	1922	5.04	1	Reb Russell	1922	4.93
2	Hack Wilson	1930	4.71	2	Cap Anson	1886	4.74
3	Sam Thompson	1894	4.62	3	Babe Ruth	1920	4.65
4	Charlie Ferguson	1887	4.61	4	Babe Ruth	1919	4.21
5	Rynie Wolters	1871	4.54	5	Babe Ruth	1921	4.21
6	Babe Ruth	1921	4.49	6	Babe Ruth	1926	4.20
7	Manny Ramirez	1999	4.47	7	Charlie Furguson	1887	4.17
8	Jimmie Foxx	1938	4.33	8	Gavvy Cravath	1913	4.05
9	Mark McGwier	2000	4.32	9	Joe Wood	1921	4.04
10	Joe Wood	1921	4.32	10	Babe Ruth	1932	4.03

Figure 1. Slugging Percentage Mean, 1871-2010, and OLS Regression on Level and Trend Breaks in 1921 and 1992

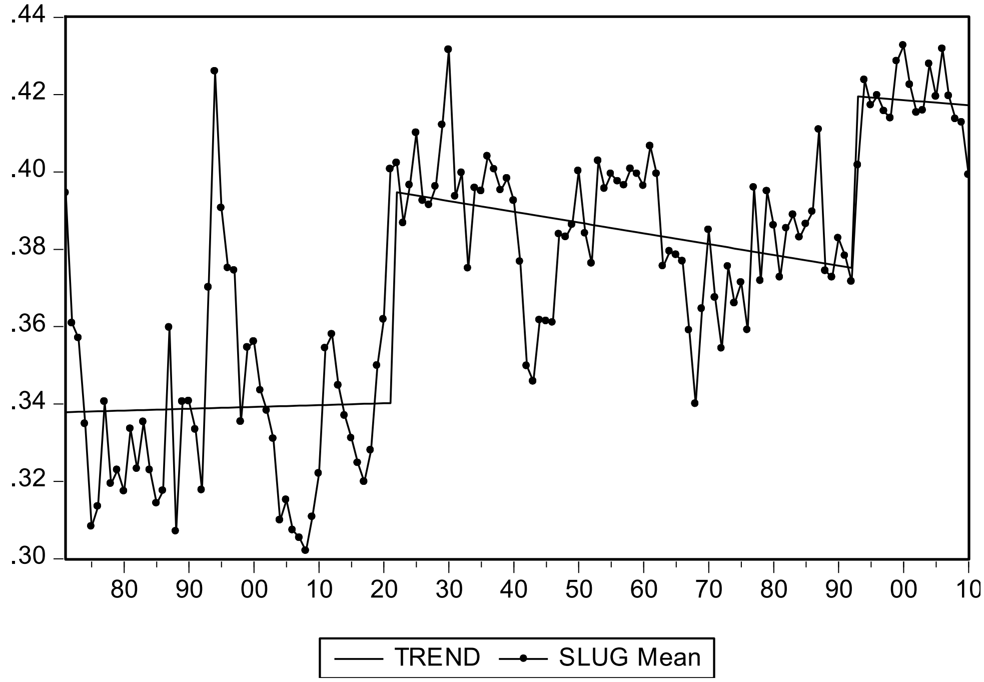

Figure 2. Home Run Standard Deviation, 1871-2010, and OLS Regression on Level and Trend Breaks in 1920 and 1966

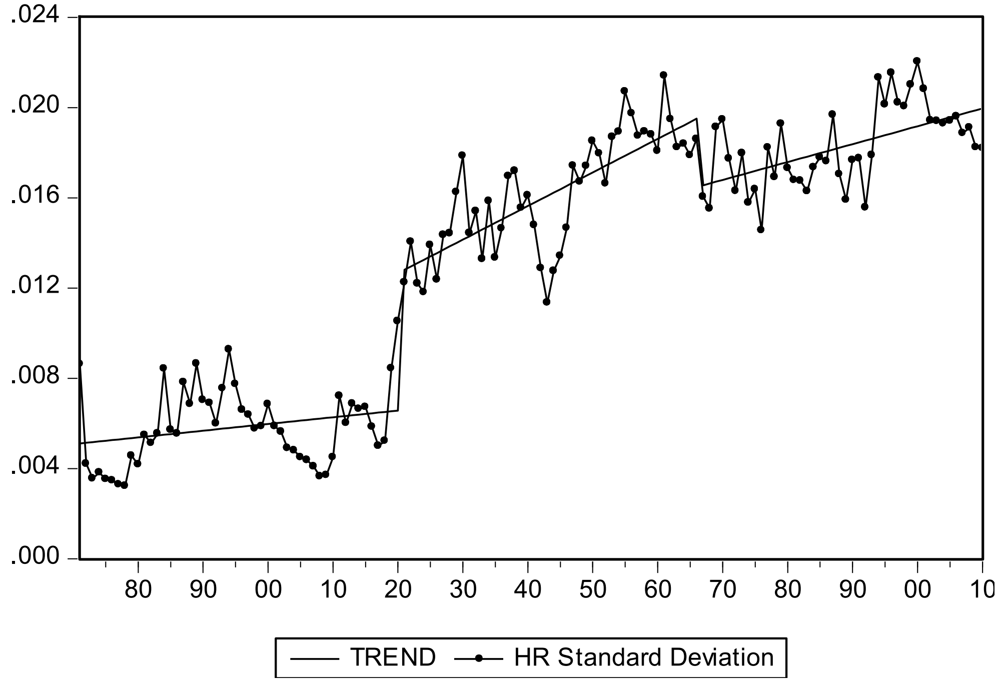

Figure 3. Batting Average Standard Deviations, 1871-2010, and OLS Regression on Level and Trend Breaks in 1906 and 1933

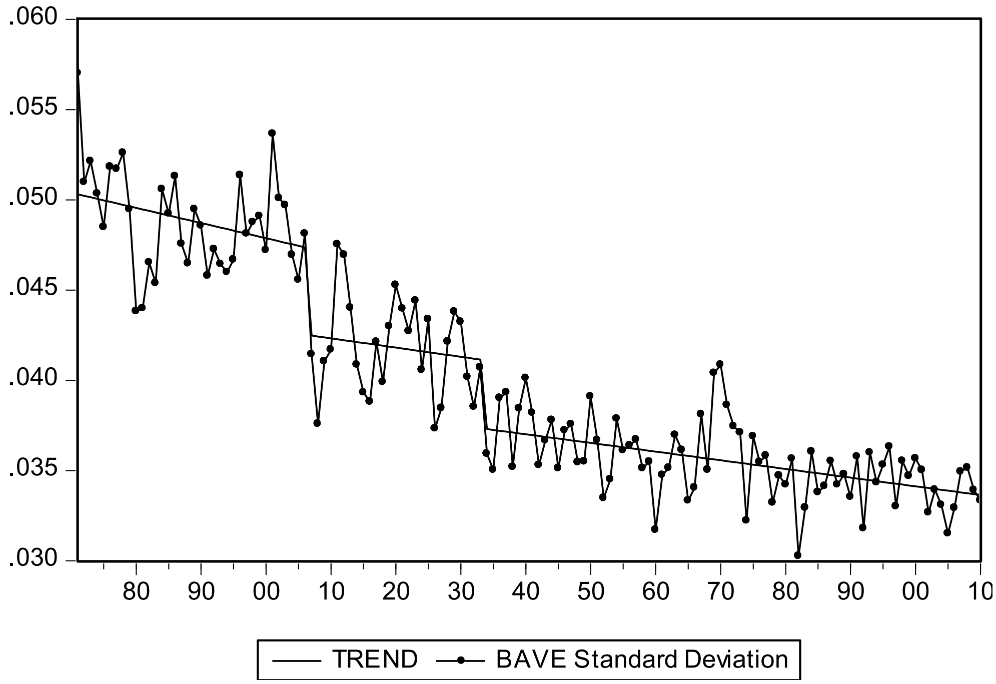

Figure 4 Runs Batted In Mean, 1871-2010, and OLS Regression on Level and Trend Break in 1887

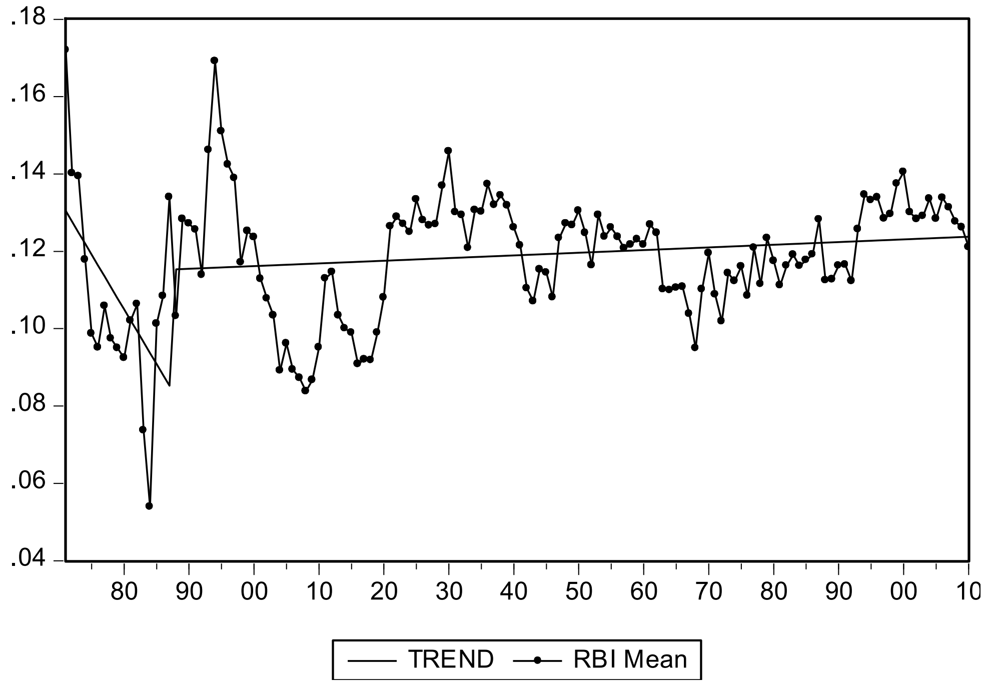

Figure 5. Runs Batted In Standard Deviation, 1871-2010, and OLS Regression on Level and Trend Break in 1921

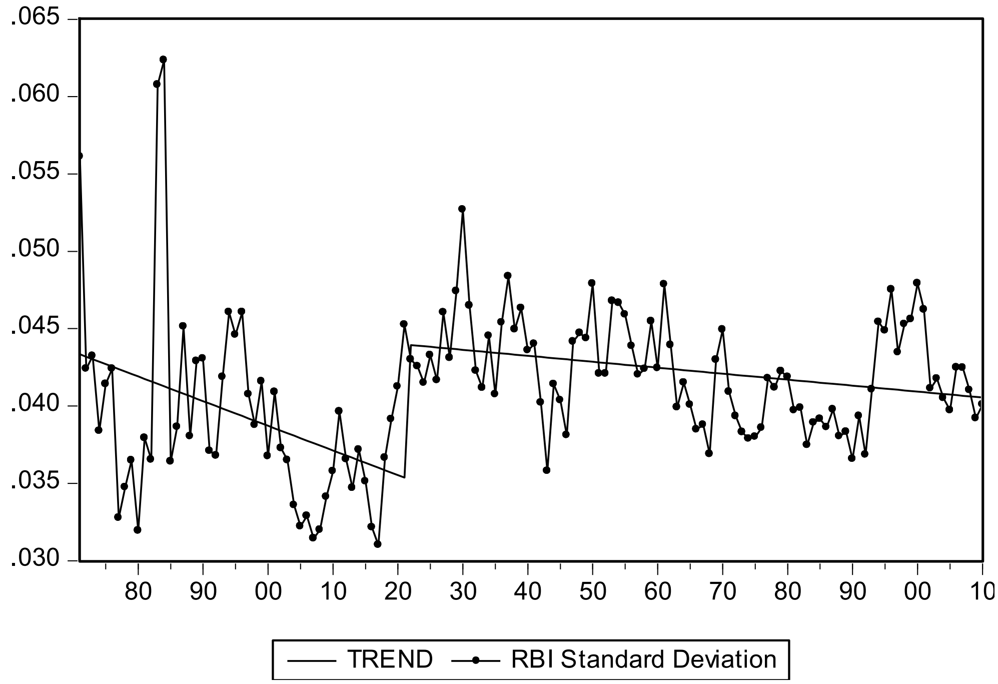

Figure 6. Slugging Percentage Standard Deviation, 1971-2010

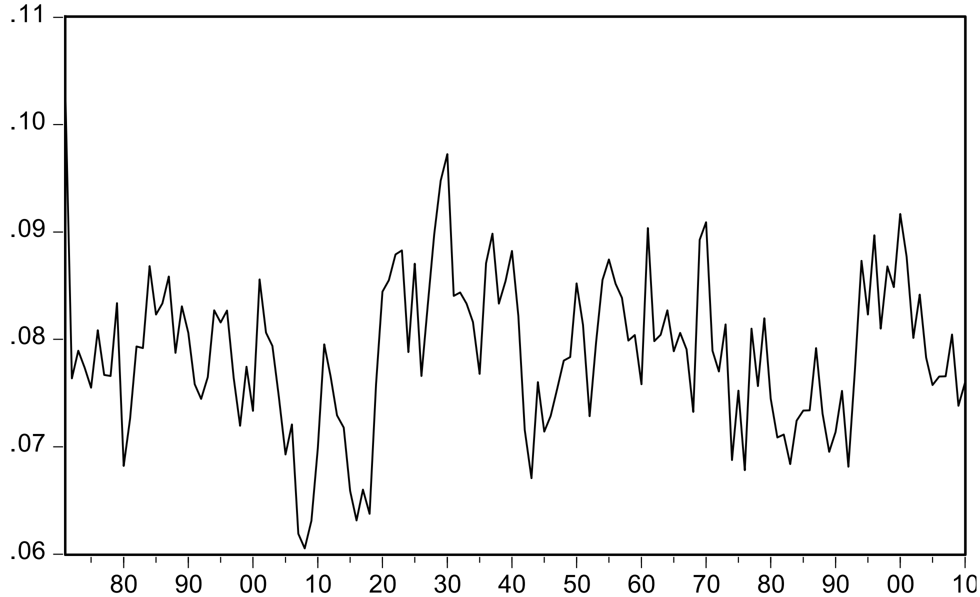

Figure 7. Home Run Mean, 1971-2010

Figure 8. Batting Average Mean, 1971-2010