Article

Journal of Sports Economics
2017, Vol. 18(6) 622-637

DOI: 10.1177/1527002515593113
journals.sagepub.com/home/jse
SAGE

Structural Breaks in the Game: The Case of Major League Baseball

Peter A. Groothuis[1], Kurt W. Rotthoff[2], and Mark C. Strazicich[3]

Abstract

To search for eras in a sports league, we utilize time-series tests with structural breaks in Major League Baseball performance. Using data from 1871-2010, the mean and standard deviation of four different performance measures are examined. Throughout, rather than assume that a break point is known a priori, we identify breaks endogenously from the data. Perhaps most notable among our findings, we identify a deterministic trend in mean slugging percentage with breaks in 1921 and 1992. Interestingly, these years closely coincide with the early years of the free-swinging (Babe Ruth) era and the modern steroid era, respectively.

Keywords

Major League Baseball, technological change, structural breaks

Introduction

Over time, games change and new innovations are developed. In many sports, the equipment drives these changes, such as innovations in tennis rackets, golf club technologies, or swimsuits (now banned). In other sports, change might result from the development of a new defensive technique (Lawrence Taylor), a new way to swing

[1] Walker College of Business, Appalachian State University, Boone, NC, USA
[2] Stillman School of Business, Seton Hall University, South Orange, NJ, USA
[3] Department of Economics, Appalachian State University, Boone, NC, USA

Corresponding Author:
Kurt W. Rotthoff, Stillman School of Business, Seton Hall University, 400 South Orange Ave, JH 674, South Orange, NJ 07079, USA.
Email: rotthoff@gmail.com

the bat (Babe Ruth), throw a pitch ("Bullet Joe" Bush), shoot a basket (Goose Tatum), or hold a putter (the current debate over the belly putter). As players adopt successful innovations, they are mimicked and the game can change.

Sports historians and scholars have often assumed *exogenous* changes in the game based on particular historical events. For instance, in baseball, one historian suggests that the game changed in 1920 when Ray Chapman was killed by a pitch that year and baseball responded by banning the spitball (Okrent, 1989). In contrast, we make no prior judgments about the timing of eras. Instead, we let the data speak to *endogenously* identify eras. To do so, we utilize time-series tests with structural breaks to identify eras.

Utilizing time-series tests to analyze sports data has become increasingly popular in the literature. For example, Davies, Downward, and Jackson (1995) and Schmidt and Berri (2004) utilize time-series tests to examine attendance in rugby and several other sports, respectively. In another line of research, Fort and Lee (2006, 2007) and Lee and Fort (2005, 2008, 2012) utilize time-series tests with structural breaks to examine competitive balance in a variety of sports. Throughout, the authors note the usefulness and importance of allowing for structural breaks when examining times series on sports data. Mills and Fort (2014) utilize similar testing procedures to examine time series on attendance. In a related line of research, Palacios-Huerta (2004) utilizes time-series tests with structural breaks to examine the mean and variance of game outcomes in soccer (football). Nieswiadomy, Strazicich, and Clayton (2012) utilize time-series tests with structural breaks to examine the batting performance of Barry Bonds. In this article, we adopt a similar methodological approach and extend the literature by utilizing time-series tests with structural breaks to examine player performance in Major League Baseball (MLB). Specifically, we examine time series on the mean and standard deviation of four traditional MLB batting performance measures. By utilizing time-series tests with breaks that are *endogenously* identified by the data, our goal is to identify eras in MLB performance that may not have been apparent when focusing a priori on particular historical events.

We find that most MLB performance measures are stationary around a deterministic trend with one or more structural breaks. Perhaps most notable among our findings, we identify a deterministic trend in the mean slugging percentage that shifts upward and changes slope in 1921 and 1992. Interestingly, these years coincide with the early years of the free-swinging (Babe Ruth) era and the modern steroid era, respectively. We conclude that structural breaks should be considered when identifying eras and comparing performance over time. Our paper proceeds as follows. The data and testing procedures are described in the second section and results in the third section. We conclude in the fourth section.

Data and Structural Breaks

MLB attracts the best baseball players in the world. The first professional baseball team was established in 1869 (the Cincinnati Red Stockings) and the league itself started in the late 1800s (and continues today). The best players in the league have their names written in the record books. However, changes in the game have led to questions about how records should be kept. For example, Barry Bonds has the most homeruns in one season, 73 in 2001, while Babe Ruth hit 60 home runs (HRs) in 1927. However, there were 162 games a season in 2001, but only 154 games in 1927 and some rules differed. Although the number of HRs per game is higher for Barry Bonds, differences in the game due to different technologies, rules, and style of play can matter.

Using data from Sean Lahman's Baseball Database on all MLB players from 1871-2010 with at least 100 at bats, we measure slugging percentage (SLUG), HRs per hundred at bats, batting average (BAVE), and runs batted in (RBI) per hundred at bats.[1] With 35,728 single-season observations, we find that the average player hit seven homeruns per season (with a maximum of 73), had 42.5 RBI, and a SLUG of .379. Using the data for each player, 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. We utilize these eight time series in our empirical investigation. A summary of the mean and standard deviation of our data is provided in Table 1.

To determine if the time-series measures of player performance are stationary around a deterministic trend or nonstationary (stochastic trend) and to look for structural breaks, we begin by utilizing the two-break minimum Lagrange Multiplier (LM) unit root test proposed by Lee and Strazicich (2003).[2] To endogenously identify the location of two breaks ($\lambda_j = T_{Bj}/T, j = 1$ and 2), the minimum LM unit root test uses a grid search to determine the combination of breaks where the unit root test statistic is minimized (i.e., the most negative). Since 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. Serial correlation is corrected by including lagged first-differenced terms selected by a sequential "general to specific" procedure.[3] The process is repeated for each combination of two breaks to jointly identify the breaks and unit root test statistic where the unit root test statistic is minimized.[4] 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), we reexamine the series using the one-break LM unit root test (Lee & Strazicich, 2013). If no break is identified, we then utilize the conventional (no-break) augmented Dickey-Fuller unit test (Dickey & Fuller, 1979, 1981). Rejection of the null indicates that the series is stationary around a deterministic trend. In contrast, failure to reject the null implies a nonstationary series with a stochastic trend.[5]

After identifying the time series that are stationary around one or two breaks, we next perform tests to see if additional breaks are present. To do so, we utilize the multiple break tests suggested by Bai and Perron (1998, 2003, BP hereafter). Given that the BP tests are valid only for stationary time series, we begin by performing regressions on the level and trend breaks for the series identified as stationary with two breaks.[6] We then apply the BP test to the residuals of these regressions to search

Table 1. Summary Statistics, Annual Major League Baseball (MLB) Performance, 1871-2010.

Variable	Mean	Median	Standard Deviation	Minimum	Maximum
SLUGM	0.372	0.377	0.034	0.302	0.433
SLUGSD	0.079	0.079	0.007	0.061	0.103
HRM	0.016	0.015	0.010	0.002	0.033
HRSD	0.013	0.015	0.006	0.003	0.022
BAVEM	0.263	0.262	0.014	0.230	0.304
BAVESD	0.040	0.038	0.006	0.030	0.057
RBIM	0.118	0.121	0.017	0.054	0.172
RBISD	0.041	0.041	0.005	0.031	0.062

Note. Mean and standard deviation of annual Major League Baseball Slugging Percentage (Slugging Percentage is calculated by total bases divided by at bats, SLUGM, and SLUGSD), Home Runs (HRM and HRSD), Batting Average (Batting Average is calculated by hits divided by at bats, batting average mean [BAVEM], and BAVESD), and Runs Batted in (RBIM and RBISD), respectively. Data are calculated from Sean Lahman's Baseball Database on all players from 1871-2010 with at least 100 at bats (http://baseball1.com/2011/01/baseball-database-updated-2010/).

for additional breaks. Given that the breaks in these regressions are the same as those that reject the unit root hypothesis in the two-break LM unit root test, the residuals of these regressions will be stationary and the BP test can be safely applied to search for additional breaks.[7]

Results

We begin by discussing the LM unit root test results displayed in Table 2. The slugging percentage mean (SLUGM) rejects the unit root null hypothesis at the 5% significance level in the two-break test, implying that SLUGM is a stationary series with structural breaks in 1921 and 1992. For the slugging percentage standard deviation (SLUGSD), only one break was significant in the two-break test. We therefore retested this series using the one-break test. In contrast to SLUGM, the SLUGSD cannot reject a unit root at the 10% level of significance, indicating that this series is nonstationary and has a stochastic trend. Similarly, the unit root hypothesis cannot be rejected for the homerun mean (HRM) at the 10% level of significance, implying that this series is nonstationary. In contrast, the unit root hypothesis is rejected for the home run standard deviation (HRSD) at the 5% level of significance, indicating that this series is stationary with structural breaks in 1920 and 1966. The batting average mean (BAVEM) cannot reject the unit root hypothesis at the 10% level of significance, implying that BAVEM is 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 breaks in 1906 and 1933. The runs batted in mean (RBIM) rejects the unit root hypothesis at the 10% level of significance, implying that RBIM is a stationary series with a

Table 2. LM Unit Root Test Results, 1871-2010.

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

Note. SLUG, HR, BAVE, and RBI denote annual slugging percentage, home runs, batting average, and runs batted in, 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 Lagrange Multiplier (LM) unit root tests come from Lee and Strazicich (2003, 2013). The critical values depend on the location of the break points, $\lambda = (T_{B1}/T, T_{B2}/T)$, and are symmetric around λ and $(1 - \lambda)$.
*, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

break in 1887. The runs batted in standard deviation (RBISD) rejects the unit root hypothesis at the 5% level of significance, implying that RBISD is a stationary series with a break in 1921.

We next perform regressions on the level and trend breaks for the five performance measures that reject the unit root hypothesis in Table 2 (SLUGM, HRSD, BAVESD, RBIM, and RBISD).[8] The results are displayed in Table 3. White's robust standard errors are used to control for heteroscedasticity. Serial correlation is corrected by including lagged values of the dependent variable identified by a similar general to specific procedure as described in Note 3.[9] In each case, Ljung-Box Q statistics for 24 lags indicate that the null hypothesis of no remaining serial correlations cannot be rejected at the 10% level of significance. Using the residuals from the regressions with two breaks in Table 3, we then apply the multiple break BP test to SLUGM, HRSD, and BAVESD to search for additional breaks.[10] A summary of the identified breaks is displayed in Table 4. Compared to the two breaks identified by the LM unit root test for SLUGM, HRSD, and BAVESD, the BP test finds two additional breaks for each of these series. For SLUGM, the BP test finds additional breaks in 1902 and 1920. Given that the break in 1920 is nearly identical to the break in 1921 identified by the LM unit root test, we will ignore this additional break in the discussion that follows. For HRSD, the BP test finds additional breaks in 1900 and 1918. Given that the break in 1918 is (again) nearly identical to the break in 1920, we will ignore this additional break in the discussion that follows. For BAVESD, the BP test finds additional breaks in 1966 and 1981.

To more carefully examine the sign and significance of including the additional breaks, we next perform regressions for the three series in Table 4 containing more

Table 3. OLS Regression Results of SLUGM, HRSD, BAVESD, RBIM, and RBISD on the Level and Trend Breaks in Table 2, 1871-2010.

$$\underset{}{\mathrm{SLUGM}_t} = \underset{(4.650)^{***}}{0.126} + \underset{(3.800)^{***}}{0.029\mathrm{D}_{1922\text{–}1992}} + \underset{(5.632)^{***}}{0.046\mathrm{D}_{1993\text{–}2010}} + \underset{(1.471)}{0.0002\mathrm{T}_{1871\text{–}1921}} - \underset{(-1.336)}{0.0001\mathrm{T}_{1922\text{–}1992}} - \underset{(-2.178)^{**}}{0.001\mathrm{T}_{1993\text{–}2010}} + \mathrm{lags}(1) + \mathrm{e}_t$$

Adjusted $R^2 = .791$ SER $= 0.016$ $Q(24) = 30.321$ (p value $= .174$)

$$\mathrm{HRSD}_t = \underset{(3.876)^{***}}{0.002} + \underset{(5.009)^{***}}{0.004\mathrm{D}_{1921\text{–}1966}} + \underset{(4.827)^{***}}{0.006\mathrm{D}_{1967\text{–}2010}} + \underset{(1.967)^{*}}{0.00003\mathrm{T}_{1871\text{–}1920}} + \underset{(3.142)^{***}}{0.0001\mathrm{T}_{1921\text{–}1966}} + \underset{(2.233)^{**}}{0.00004\mathrm{T}_{1967\text{–}2010}} + \mathrm{lags}(1) + \mathrm{e}_t$$

Adjusted $R^2 = .947$ SER $= 0.001$ $Q(24) = 31.507$ (p value $= .140$)

$$\mathrm{BAVESD}_t = \underset{(7.740)^{***}}{0.051} - \underset{(-4.435)^{***}}{0.006\mathrm{D}_{1907\text{–}1933}} - \underset{(-6.502)^{***}}{0.012\mathrm{D}_{1934\text{–}2010}} - \underset{(-0.598)}{0.00003\mathrm{T}_{1871\text{–}1906}} - \underset{(-1.186)}{0.0001\mathrm{T}_{1907\text{–}1933}} - \underset{(-4.674)^{***}}{0.0001\mathrm{T}_{1934\text{–}2010}} + \mathrm{lags}(4) + \mathrm{e}_t$$

Adjusted $R^2 = .875$ SER $= 0.002$ $Q(24) = 25.240$ (p value $= .393$)

$$\mathrm{RBIM}_t = \underset{(0.263)}{0.005} + \underset{(1.245)}{0.018\mathrm{D}_{1888\text{–}2010}} + \underset{(0.989)}{0.002\mathrm{T}_{1871\text{–}1987}} + \underset{(0.688)}{0.00002\mathrm{T}_{1888\text{–}2010}} + \mathrm{lags}(5) + \mathrm{e}_t$$

Adjusted $R^2 = .639$ SER $= 0.010$ $Q(24) = 17.295$ (p value $= .836$)

$$\mathrm{RBISD}_t = \underset{(2.930)^{***}}{0.025} + \underset{(0.185)}{0.0005\mathrm{D}_{1922\text{–}2010}} - \underset{(-1.040)}{0.00008\mathrm{T}_{1871\text{–}1921}} - \underset{(-1.635)}{0.00002\mathrm{T}_{1922\text{–}2010}} + \mathrm{lags}(4) + \mathrm{e}_t$$

Adjusted $R^2 = .396$ SER $= 0.004$ $Q(24) = 14.026$ (p value $= .946$)

Note. Dependent variable is the slugging percentage mean (SLUGM), home runs standard deviation (HRSD), batting average standard deviation (BAVESD), runs batted in mean (RBIM), and runs batted in standard deviation (RBISD), in year t, respectively. t-Statistics are shown in parentheses. D and T represent dummy variables for the intercept and trend breaks identified using the Lagrange Multiplier (LM) unit root test as reported in Table 2. White's robust standard errors were utilized to control for heteroscedasticity. Lagged values of the dependent variable were included to correct for serial correlation using the method described in Note 3. The Ljung-Box Q statistic for 24 lags tests the null of no remaining serial correlations in the residuals. SER denotes the standard error of the regression.
***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

Table 4. Summary of Structural Breaks in the Stationary Series, 1871-2010.

Time Series	LM Test Breaks	BP Test Breaks
SLUGM	1921, 1992	1902, 1920
HRSD	1920, 1966	1900, 1918
BAVESD	1906, 1933	1966, 1981
RBIM	1887	–
RBISD	1921	–

Note. SLUG, HR, BAVE, and RBI denote annual slugging percentage, home runs, batting average, and runs batted in of all players in the series, where *M* denotes the mean and *SD* denotes the standard deviation, respectively. The Lagrange Multiplier (LM) test breaks are those identified with the Lee and Strazicich (2003, 2013) one- and two-break LM unit root tests. The Bai and Perron (BP) test breaks are the breaks identified by applying the Bai and Perron (1998, 2003) procedure to the stationary residuals from the regressions with two level and trend breaks in Table 3.

than two breaks (i.e., SLUGM, HRSD, and BAVESD). Robust standard errors and lagged dependent variables are again included in each regression to control for heteroscedasticity and serial correlation. The results are reported in Table 5.[11] We begin by discussing results for the slugging percentage means (SLUGM). There is a downward shift in SLUGM in 1902, the additional break identified by the BP test, while the break coefficient is not statistically significant. Following this, there is a significant upward shift in the mean slugging percentage in 1921 and another significant upward shift in 1992. The break in 1921 coincides with the early years of the free-swinging (Babe Ruth) era while 1992 coincides with the early years of the modern steroid era often associated with Jose Canseco and Mark McGwire, among others.[12] Following the upward shift in 1992, there is a small but significant downward trend in SLUGM.

We next examine regression results for the HRSD. Following the break in 1900, there is a small upward shift in HRSD, while the break is not statistically significant. Following the break in 1920, there is a significant upward shift in HRSD indicating that the dispersion in HR performance increased. After the final break in 1966, there is a significant upward shift in HRSD with a small but significant upward trend. Perhaps most notable among these findings is that we again see a significant structural break in 1920 associated with the early years of the free-swinging (Babe Ruth) era. These findings provide additional support to the idea that the Babe Ruth era had a major influence on the game. During this time, there were two major events: Babe Ruth influenced the game by creating a batting style that could be, and was, emulated by other players. Moreover, at the same time, the type of ball and style of pitching changed. We suggest that combined these changes had a significant impact on the game, which led to a greater dispersion in HR performance among players.

We next examine the regression results for the BAVESD. Following the break in 1906, there is a significant downward shift in BAVESD followed by further significant downward shifts in 1933, 1966, and 1981.

Table 5. OLS Regression Results of SLUGM, HRSD, and BAVESD on the Level and Trend Breaks in Table 4, 1871-2010.

$$
\begin{aligned}
\mathrm{SLUGM}_t = \; & \underset{(5.817)^{***}}{0.147} - \underset{(-1.235)}{0.01\mathrm{D}_{1903\text{–}1921}} + \underset{(5.034)^{***}}{0.038\mathrm{D}_{1922\text{–}1992}} + \underset{(7.072)^{***}}{0.057\mathrm{D}_{1993\text{–}2010}} + \underset{(2.111)^{**}}{0.001\mathrm{T}_{1871\text{–}1902}} + \underset{(3.132)^{***}}{0.002\mathrm{T}_{1903\text{–}1921}} \\
& \underset{(-1.612)}{-0.0001\mathrm{T}_{1922\text{–}1992}} \underset{(-2.002)^{**}}{-0.001\mathrm{T}_{1993\text{–}2010}} + \mathrm{lags}(1) + \mathrm{e}_t
\end{aligned}
$$

Adjusted $R^2 = .806$ SER $= 0.015$ $Q(24) = 29.077$ (p value $= .217$)

$$
\begin{aligned}
\mathrm{HRSD}_t = \; & \underset{(3.630)^{***}}{0.002} + \underset{(0.190)}{0.0001\mathrm{D}_{1901\text{–}1920}} + \underset{(5.832)^{***}}{0.005\mathrm{D}_{1921\text{–}1966}} + \underset{(5.539)^{***}}{0.007\mathrm{D}_{1967\text{–}2010}} + \underset{(3.594)^{***}}{0.00009\mathrm{T}_{1871\text{–}1900}} \\
& + \underset{(2.638)^{***}}{0.0001\mathrm{T}_{1901\text{–}1920}} + \underset{(3.579)^{***}}{0.00007\mathrm{T}_{1921\text{–}1966}} + \underset{(2.531)^{**}}{0.00004\mathrm{T}_{1967\text{–}2010}} + \mathrm{lags}(1) + \mathrm{e}_t
\end{aligned}
$$

Adjusted $R^2 = .950$ SER $= 0.001$ $Q(24) = 33.968$ (p value $= .085$)

$$
\begin{aligned}
\mathrm{BAVESD}_t = \; & \underset{(8.760)^{***}}{0.059} - \underset{(-5.093)^{***}}{0.007\mathrm{D}_{1907\text{–}1933}} - \underset{(-7.161)^{***}}{0.013\mathrm{D}_{1934\text{–}1966}} - \underset{(-5.929)^{***}}{0.013\mathrm{D}_{1967\text{–}1981}} - \underset{(-7.595)^{***}}{0.018\mathrm{D}_{1982\text{–}2010}} - \underset{(-0.729)}{0.00004\mathrm{T}_{1871\text{–}1906}} \\
& \underset{(-1.573)}{-0.00009\mathrm{T}_{1907\text{–}1933}} - \underset{(-3.605)^{***}}{0.0001\mathrm{T}_{1934\text{–}1966}} - \underset{(-2.479)^{**}}{0.0003\mathrm{T}_{1967\text{–}1981}} - \underset{(-0.109)}{0.000004\mathrm{T}_{1982\text{–}2010}} + \mathrm{lags}(4) + \mathrm{e}_t
\end{aligned}
$$

Adjusted $R^2 = .884$ SER $= 0.002$ $Q(24) = 25.611$ (p value $= .373$)

Note. Dependent variable is the slugging percentage mean (SLUGM), home runs standard deviation (HRSD), and batting average standard deviation (BAVESD) in year *t*, respectively. *t*-Statistics are shown in parentheses. *D* and *T* represent dummy variables for the intercept and trend breaks reported in Table 4. White's robust standard errors were utilized to control for heteroscedasticity. Lagged values of the dependent variable were included to correct for serial correlation using the method described in Note 3. The Ljung-Box *Q* statistic for 24 lags tests the null of no remaining serial correlations in the residuals.

***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

Next, we examine results for the mean and standard deviation of RBI (RBIM and RBISD) using the single break regression results in Table 3. Following the break in 1887, there is an upward shift in RBIM. Although there is a slight positive trend following the break, the trend slope is not significant (at the 10% level). For RBISD, following the break in 1921 there is an upward shift in RBISD. There is a slight negative trend following the break, but the trend slope is not significant (at the 10% level). Although the break in RBISD is again associated with the Babe Ruth era, there is little notable change in the series before and after the break. Overall, the results for RBI suggest that there has been little change in RBIM and RBISD.

To better visualize the regression results of the stationary series with breaks reported in Table 5 (SLUGM, HRSD, and BAVESD) and Table 3 (RBIM and RBISD), we next construct simple plots of the estimated trends and actual data in Figures 1–5. As in the regression results described previously, perhaps most interesting is the plot of the slugging percentage means (SLUGM) displayed in Figure 1. From Figure 1, we can clearly observe the significant upward shifts in SLUGM that occurred in 1921 and 1992. Similarly, in Figure 2, we see a notable upward shift in the HRSD in 1920. In Figure 3, we can observe the general decline in the BAVESD. In Figures 4 and 5, we observe the relative stability of the mean and standard deviation of RBI (RBIM and RBISD), respectively.

The above-mentioned results suggest that MLB performance had a notable structural break around 1920-1921. The trend in the slugging percentage mean (SLUGM) had a significant upward shift in 1921, suggesting that the average player began to increasingly hit for power and hit more doubles, triples, and HRs. Similarly, the trend in the HRSD had a significant upward shift and increasing slope in 1920 and the RBISD had a positive (small) upward shift in 1921. These findings suggest that players after 1920 became more diverse in their performance with some players increasingly hitting for power while others did not. Combined with the 1921 break in the slugging percentage mean (SLUGM), these findings suggest the possibility that following the success of Babe Ruth's free-swinging style, combined with the required pitching changes that occurred at this time, others could have mimicked his innovation and hit for power as well.

It is interesting to note that our *endogenous* structural break in 1921 occurs close to a break suggested by the baseball historian Stanley Rothman. To provide historical context to this era, Rothman (2012) suggests that the 1920s changed the offensive strategies in baseball due to three major events: The Black Sox scandal, the banning of the spitball, and Babe Ruth's style of hitting being copied. The Black Sox scandal occurred in 1919 and caused the popularity of baseball to decline. Rothman further suggests that in an attempt to clean up the Black Sox scandal of 1919, the spitball and other trick pitches were abolished starting in 1920 (which also came after the death of Ray Chapman in 1920, for which "umpires were ordered to only use shiny white balls throughout the game"). Rothman further suggests that in 1919, when Babe Ruth hit a record 29 HRs during his last year with the Red Sox, that this

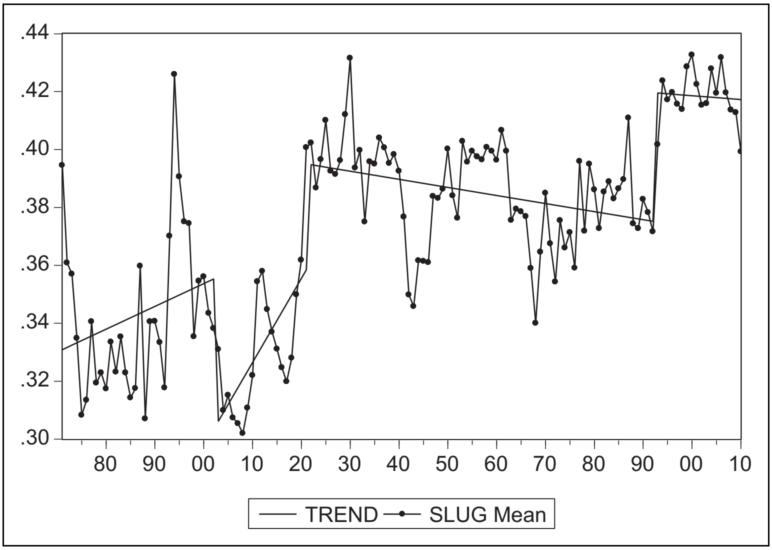

Figure 1. Slugging percentage mean, 1871-2010, and ordinary least square (OLS) regression on level and trend breaks in 1902, 1921, and 1992.

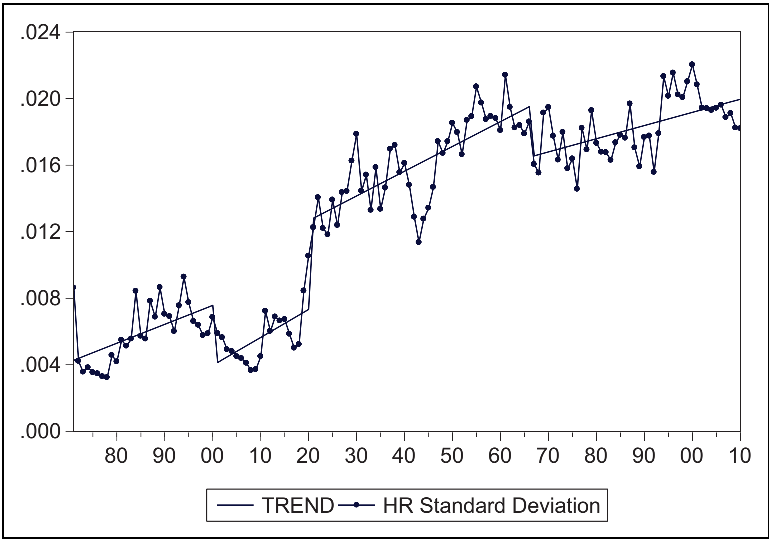

Figure 2. Home run standard deviation (HRSD), 1871-2010, and ordinary least square (OLS) regression on level and trend breaks in 1900, 1920, and 1966.

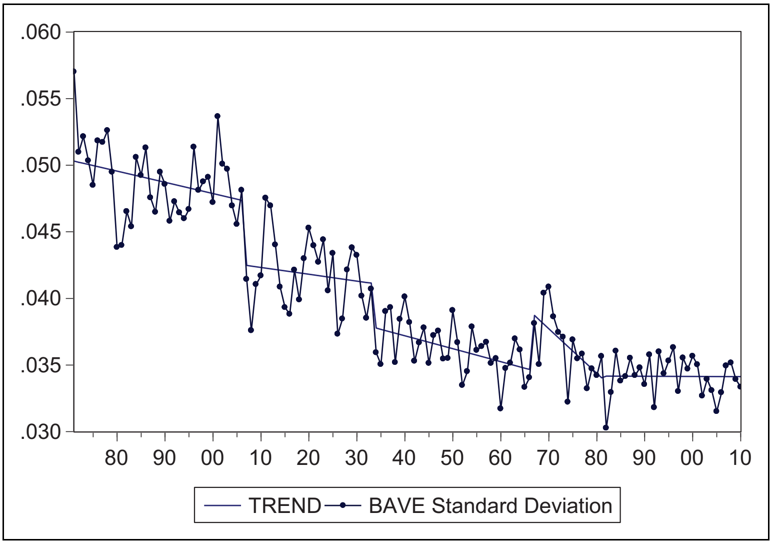

Figure 3. Batting average standard deviations (BAVESD), 1871-2010, and ordinary least square (OLS) regression on level and trend breaks in 1906, 1933, 1966, and 1981.

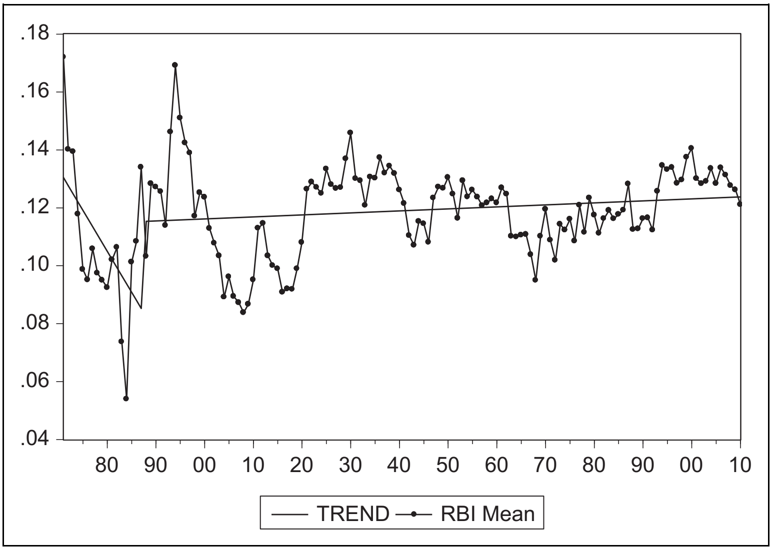

Figure 4. Runs batted in mean (RBIM), 1871-2010, and ordinary least square (OLS) regression on level and trend break in 1887.

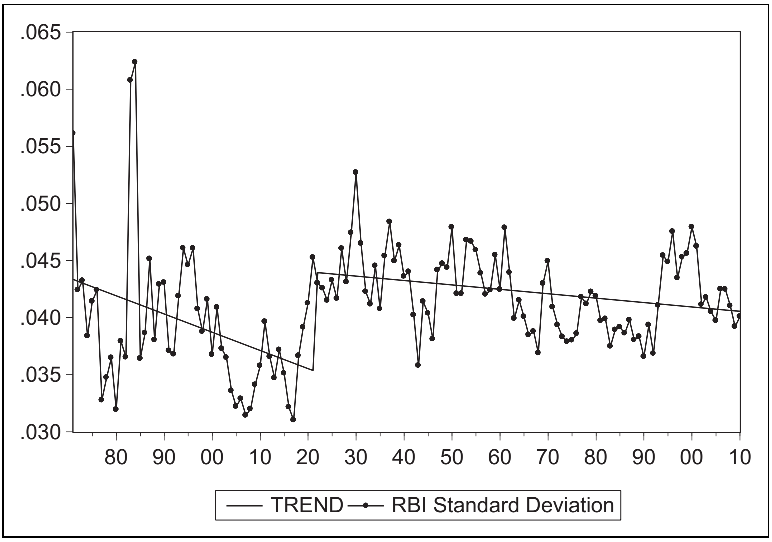

Figure 5. Runs batted in standard deviation (RBISD), 1871-2010, and ordinary least square (OLS) regression on level and trend break in 1921.

created a notable increase in fan excitement about HRs. These events coincide with our suggestion that starting in 1920 or 1921 many players increasingly copied Babe Ruth's hitting style. This style included not choking up on the bat and swinging with a pronounced uppercut. Ruth's style replaced the spray-hitting style that resulted in many singles. Rothman also notes that attendance in 1920 increased approximately 20% from 1919. As owners noticed this increase in attendance, they attributed this to an increase in offensive performance and HR hitting, which encouraged an increased adoption of this new style of hitting.

We suggest that the structural break that we identify in 1921 is consistent with all three of the events described above, during what most people remember as the "Babe Ruth era." Two additional events also occurred during this time period. First, Lee and Fort (2008) find a structural break in 1918 associated with an increase in MLB attendance after the end of World War I. Second, this time period is also associated with the advent of the professional Negro National league in 1920. We suggest that in the pursuit of greater attendance at MLB games, these additional events may also have contributed to a change in the game.

Most notable among our other breaks is the significant upward shift in the slugging percentage mean (SLUGM) in 1992, which is closely associated with (what many perceive to be) the early years of the modern steroid era. Although it is difficult to identify the start of the steroid era since players attempt to keep hidden the use of performance enhancing drugs, it is interesting to note that while MLB banned steroids in 1991 they did not test for their usage until 2003. We suggest that fan interest in offensive performance (Ahn & Lee, 2014), coupled with the knowledge that steroids were becoming an issue in MLB due to the explicit ban, might have caused the structural break in 1992 as players mimicked steroid use knowing they would not be caught.

Conclusion

Over time, games change. Most analysts identify *exogenous* changes in the game based on particular historical events. In contrast, we utilize time-series techniques to *endogenously* determine where changes occur. Using several time series on MLB performance from 1871-2010 (140 seasonal observations for each series), we attempt to determine if structural breaks can be used to identify eras.

Using several measures of batting performance, we perform time-series tests to identify whether deterministic or stochastic trends and structural breaks are present. Perhaps most notable among our findings, we identify significant structural breaks in 1921 and 1992 for the mean slugging percentage. Given that the breaks are endogenously determined from the data, it is interesting to analyze what was happening in the sport during these years.

Interestingly, the break in 1921 coincides with the early years of the free-swinging era often identified with Babe Ruth. We suggest that during this time period, other players began to mimic Babe Ruth's free-swinging style. In addition, pitching rules changed at this time. Combined, we suggest that these events had a major influence on baseball and changed the game. Finally, our identified break in 1992 is consistent with historical observations commonly associated with the early years of the modern steroid era. In particular, it is interesting to note that 1992 is 1 year after MLB enacted a ban on steroid use while the ban was not enforced with testing until 2003.

In conclusion, we suggest that the findings presented here support utilizing time-series tests with structural breaks to identify eras in sports and provide insights for future research. Given that our current analysis has examined hitting statistics, we suggest that future research should focus on time series of pitching and league-level performance data to search for additional insights into changes in the game.

Acknowledgment

We thank the editor, three anonymous referees, and seminar participants at Appalachian State University and the Eastern Economic Association annual meeting in 2012.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

Notes

1. 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).
2. 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.
3. To determine the number of lagged first-differenced terms that correct for serial correlation, we employ the following sequential "general to specific" procedure. At each combination of two break points over the time interval [.1 T, .9 T] (to eliminate end points), we begin with a maximum number of $k = 8$ lagged first-difference 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 reestimated using $k = 7$ terms, and so on until the maximum lagged term is found (i.e., the order of serial correlation is identified and corrected), or $k = 0$ (i.e., there is no serial correlation). Once the maximum number of lagged terms is found, all lower lags remain in the unit root test. This type of procedure has been shown to perform well compared to other data-dependent procedures to select the optimal k and correct for serial correlation (e.g., Ng & Perron, 1995).
4. Gauss codes for the one- and two-break minimum Lagrange Multiplier (LM) unit root test are available from the authors upon request.
5. Note that the interpretation of breaks in a nonstationary (unit root) series differs from that in a stationary series. In a unit root process, a structural break in the level can be interpreted as an unusually large onetime shock or outlier, while a break in the trend can be interpreted as a permanent change in the drift.
6. See Prodan (2008) for a discussion of pitfalls that can arise when applying the Bai and Perron (1998, 2003) type tests to nonstationary time series.
7. See Lee and Fort (2005, 2008, 2012), Fort and Lee (2006, 2007), Bai and Perron (1998, 2003), and Lee and Strazicich (2003, 2013) for more detailed discussion of the tests described previously.
8. Regressions will not be undertaken for SLUGSD, HRM, and BAVEM, since these time series are nonstationary and spurious regressions can occur.
9. See Ashley (2012) for discussion of why modeling serial correlation with lagged variables is a desirable procedure.
10. When searching for breaks we impose a minimum of 15 time periods between them as suggested in Bai and Perron (2003).
11. The Ljung-Box Q-statistics for 24 lags indicate that the null hypothesis of no remaining serial correlations cannot be rejected at the 10% level of significance for SLUGM and BAVESD and at the 5% level for HRSD.

12. For example, in a statement released on the website of the St. Louis Cardinals on January 11, 2010, Mark McGwire admitted using performance-enhancing drugs beginning in 1989 or 1990.

References

Ahn, S. C., & Lee, Y. H. (2014). Major league baseball attendance: Long-term analysis using factor models. *Journal of Sports Economics*, *15*, 451–477.

Ashley, R. A. (2012). *Fundamentals of applied econometrics*. Hoboken, NJ: John Wiley.

Bai, J., & Perron, P. (1998). Estimating and testing linear models with multiple structural changes. *Econometrica*, *66*, 47–78.

Bai, J., & Perron, P. (2003). Computation and analysis of multiple structural change models. *Journal of Applied Econometrics*, *18*, 1–22.

Davies, B., Downward, P., & Jackson, I. (1995). The demand for rugby league: Evidence from causality tests. *Applied Economics*, *27*, 1003–1007.

Dickey, D. A., & Fuller, W. A. (1979). Distribution of the estimator for autoregressive time series with a unit root. *Journal of the American Statistical Association*, *74*, 427–431.

Dickey, D. A., & Fuller, W. A. (1981). Likelihood ratio statistics for autoregressive time series with a unit root. *Econometrica*, *49*, 1057–1072.

Enders, W. (2010). *Applied econometric time series* (3rd ed.). Hoboken, NJ: John Wiley & Sons, Inc.

Fort, R., & Lee, Y. H. (2006). Stationarity and major league baseball attendance analysis. *Journal of Sports Economics*, *7*, 408–415.

Fort, R., & Lee, Y. H. (2007). Structural change, competitive balance, and the rest of the major leagues. *Economic Inquiry*, *45*, 519–532.

Lee, Y. H., & Fort, R. (2005). Structural change in baseball's competitive balance: The great depression, team location, and racial integration. *Economic Inquiry*, *43*, 158–169.

Lee, Y. H., & Fort, R. (2008). Attendance and the uncertainty-of-outcome hypothesis in baseball. *Review of Industrial Organization*, *33*, 281–295.

Lee, Y. H., & Fort, R. (2012). Competitive balance: Time series lessons from the English premier league. *Scottish Journal of Political Economy*, *59*, 266–282.

Lee, J., & Strazicich, M. C. (2003). Minimum Lagrange multiplier unit root test with two structural breaks. *Review of Economics and Statistics*, *8*, 1082–1089.

Lee, J., & Strazicich, M. C. (2013). Minimum LM unit root test with one structural break. *Economics Bulletin*, *33*, 2483–2492.

Mills, B., & Fort, R. (2014). League-level attendance and outcome uncertainty in U.S. pro sports leagues. *Economic Inquiry*, *52*, 205–218.

Nieswiadomy, M. L., Strazicich, M. C., & Clayton, S. (2012). Was there a structural break in Barry Bonds's Bat? *Journal of Quantitative Analysis in Sports*, *8*(3), ISSN 1559-0410, DOI: 10.1515/1559-0410.1305, October 2012.

Ng, S., & Perron, P. (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*, 269–281.

Okrent, D. (1989). *Baseball anecdotes*. Oxford, England: Oxford University Press.
Palacios-Huerta, I. (2004). Structural changes during a Century of the World's most popular sport. *Statistical Methods and Applications*, *13*, 241–258.
Prodan, R. (2008). Potential pitfalls in determining multiple structural changes with an application to purchasing power parity. *Journal of Business and Economic Statistics*, *26*, 50–65.
Rothman, S. (2012). *Sandlot stats: Learning statistics with baseball*. Baltimore, MD: The John Hopkins University Press.
Schmidt, M. B., & Berri, D. J. (2004). The impact of labor strikes on consumer demand: An application to professional sports. *American Economic Review*, *94*, 344–357.

Author Biographies

Peter A. Groothuis is a professor in the Walker College of Business at Appalachian State University. His research interests include sports economics, labor economics, and environmental economics.

Kurt W. Rotthoff is an associate professor of economics and finance at the Stillman School of Business at Seton Hall University. His research interests are financial economics, industrial organization, economics of education, and economics and finance in sports.

Mark C. Strazicich is a professor of Economics in the Walker College of Business at Appalachian State University. His research interests include macroeconomics, economic growth and the environment, and applying time series tests to a variety of areas.