Was There a Structural Break in Barry Bondsâ Bat?

Article *in* Journal of Quantitative Analysis in Sports · January 2012

DOI: 10.1515/1559-0410.1305 · Source: RePEc

3 authors, including:

Michael Nieswiadomy
University of North Texas
62 PUBLICATIONS **1,594** CITATIONS

SEE PROFILE

Mark Strazicich
Appalachian State University
54 PUBLICATIONS **5,951** CITATIONS

SEE PROFILE

Appalachian
STATE UNIVERSITY

Department of Economics Working Paper

Number 10-13 | August 2010

Was There a Structural Break in Barry Bonds' Bat?

Stephen Clayton
University of North Texas

Michael Nieswiadomy
University of North Texas

Mark C. Strazicich
Appalachian State University

Department of Economics
Appalachian State University
Boone, NC 28608
Phone: (828) 262-6123
Fax: (828) 262-6105
www.business.appstate.edu/economics

Was There a Structural Break in Barry Bonds' Bat?

Stephen Clayton[a]

Michael Nieswiadomy[a*]

Mark C. Strazicich[b]

August 31, 2010

Keywords: age-effects, peak performance, baseball, OPS, structural break

JEL Classifications: J24, C22

[a]Department of Economics

1155 Union Circle Box 311457

University of North Texas

Denton, Texas 76201-5017

Phone: 940-565-2244

Fax: 940-565-4426

Email: stephen.clayton@unt.edu

Email: miken@unt.edu

*Corresponding author.

[b]Department of Economics

Appalachian State University

Boone, North Carolina 28608

Phone: 828-262-6124

Fax: 828-262-6105

Email: strazicichmc@appstate.edu

Was There a Structural Break in Barry Bonds' Bat?

August 31, 2010

Abstract

We utilize time series tests to investigate if Barry Bonds' batting has a deterministic or stochastic trend and to test if structural breaks occur. Bonds' monthly on base percentage plus slugging percentage (OPS) is examined from 1986 to 2007. We find that Bonds' OPS is stationary around two level and trend breaks. We find that Bonds' OPS initially follows a positive trend to the age of 28.9 (June 1993), which coincides roughly with the expected peak performance age (27.6) for a MLB batter as identified by Fair (2008). Following this break, we find that Bonds' OPS was on a plateau until a second break in September 2000. At this point, at the age of 36.1, Bonds' OPS jumps up unexpectedly and declines slowly thereafter until his retirement in September 2007 at age 43.

Keywords: age-effects, peak performance, baseball, OPS, structural break

JEL Classifications: J24, C22

I. Introduction

In this paper, we utilize time series unit root tests to investigate if Barry Bonds' batting productivity has a deterministic or stochastic trend and if structural breaks are present. To do so, we examine monthly data on his on base percentage plus slugging percentage (OPS) over the time period 1986 to 2007. We find that Bonds' OPS is characterized as a stationary process with two level and trend breaks occurring in June 1993 and September 2000. The first break occurred at age 28.9 (June 1993), which is roughly the expected age (27.6) for peak batting performance in Major League Baseball (MLB) (Fair, 2008). The second break is unexpected, occurring at age 36.1 years in September 2000. Moreover, following the break in September 2000, Bonds' batting performance increased dramatically and, while declining thereafter, remained high until his retirement in September 2007 at the age of 43. While there is no publicly known proof that Barry Bonds took performance enhancing drugs (PEDs), our findings are consistent with the timing of the 'training and products' provided by Anderson and BALCO.

Our paper proceeds as follows. Section II discusses some previous studies on hitting production and age effects in sports. Section III examines the specifics of Barry Bonds' career and allegations of PED usage with the purpose of providing a background to understand his lifetime production. Section IV presents the data. Section V describes the Lagrange Multiplier (LM) structural break unit root tests used for analysis of the data and discusses the results. Section VI concludes.

II. Baseball's Hitting Productivity Studies

Across many sports, it has been well documented that as a player's age increases his/her skill and productivity first increase, reaching a maximum after some experience has been

accumulated, and then decline as age-effects overpower further gains from experience (Scully, 1989; Sommers, 2008; Fair, 1994, 2007, 2008). While peaks differ between sports, a player's performance generally increases at a decreasing rate to its maximum and declines thereafter. In baseball, the peak performance typically occurs between the sixth and seventh years (Scully, 1989) or around 27.6 years of age (Fair, 2008).

Fair (2008) models the progression of batters in MLB in an asymmetric manner by using two conjoined quadratic curves that meet at a single peak. Taken together, the conjoined curves show the changes in performance of a MLB batter with respect to age and predict a peak in production at age 27.6. Performance is measured by "on base percentage plus slugging" (OPS), which measures the efficiency of a batter's hitting by combining on base percentage with slugging percentage to evaluate plate discipline and power. In Fair's (2008) model, each player is held to the same two-part quadratic form, but is assigned an individual constant that determines two unique intercepts and shifts the potential peak in performance up or down.

Fair (2008) examined the 441 batters who played at least 10 full seasons between 1921 and 2004 in MLB. Of the 441 batters, 423 had performances that followed the predicted pattern. Only 18 batters had results that deviated from the quadratic curves of his model, none more starkly than Barry Bonds. In our paper, we utilize Fair's (2008) model in conjunction with Barry Bond's data to calculate first a monthly predicted OPS and second a OPS residual that compares the predicted OPS to the actual OPS. Fair (2008) notes that of the 18 players whose output differed significantly from the predicted values, 15 played after 1990, within the so-called Steroid Era in baseball. Fair's (2008) equations are as follows:

$$y_{it} = \begin{cases} \alpha_{1i} + \beta_1 x_{it} + \gamma_1 x_{it}^2 + \epsilon_{it}, & x_{it} \leq \delta \\ \alpha_{2i} + \beta_2 x_{it} + \gamma_2 x_{it}^2 + \epsilon_{it}, & x_{it} \geq \delta \end{cases} \quad (1)$$

$$\beta_1 = -2\gamma_1 \delta \quad (2)$$

$$\beta_2 = -2\gamma_2$$

$$\alpha_{2i} = \alpha_{1i} + (\gamma_2 - \gamma_1)\delta^2,$$

where y_{it} is the OPS for player i in year t, x_{it} is the age of player i in year t, δ is the peak performance age, α_{1i} is the intercept for a player younger than 27.6 (δ), α_{2i} is the intercept for a player older than 27.6 (δ) and ε_{it} is an *iid* random error. Part one of Equation (1) is used to calculate the OPS of a player before his peak performance age of 27.6 years and part two is used to calculate the OPS of a player after his peak performance age, as constrained by equation (2). Together, equations (1) and (2) create two quadratic curves with different curvatures that meet at a peak at age 27.6 (δ).

Sommers (2008) examines the changing hitting profile of MLB players from 1966 to 2006 with respect to batting average and concludes that players in 1976, 1986 and 1996 all peaked later in life, at a higher maximum and with less variation in their total growth than the players in 1966. Sommers (2008) further notes that in 2006 there was a precipitous drop in the production of American League players' output in terms of batting average and speculates that this decline could be related to MLB's introduction of a PED policy in 2004. Prior to 2004, baseball had no official policy prohibiting the use of PEDs and speculation has abounded as to the pervasiveness of their use. The spread of steroid use in sports has recently been explored as a game theory exercise and the conclusion in many cases is that as long as players believe that other players are using steroids, rational players will use them (Haugen, 2004; Eber, 2008; Leeds and von Allmen, 2007). Whether adopting a standard two player model (Haugen, 2004), a two player model with a fair play consideration (Eber, 2008), or a three player model (Leeds and von Allmen, 2008; Harrington, 2009), if one player is assumed to be using steroids, others will reciprocate to avoid falling behind. This is the situation speculated about Bonds. It is noted that

Bonds may have felt compelled to use PEDs because he felt that the accomplishments of cheating players would overshadow his own accomplishments (Fainaru-Wada and Williams, 2006).

Speculation about the surge in prodigious home run hitting in recent years by Barry Bonds, Mark McGwire, and others has led to some question if these events are anomalies and if steroid use may be the cause. Devany (2010) argues that they are not anomalies, but can be explained by power laws. Dinardo and Winfree (2010) disagree; they believe that the power law explanation is flawed and it is possible (but not proven) that a judicious use of steroids could bolster a batter's home run output. Schmotzer, Switchenko, and Kilgo (2008) find evidence that steroid use increased hitting productivity 12.6% during the 1995-2007 period. Addona and Roth (2010) found that steroid use improved baseball pitchers fastball speed by an average of 1.074 MPH for the 2002 to 2008 period. The controversy surrounding possible steroid usage continues to grow, even prompting investigations by Congress and federal grand juries (ESPN, 2010). In August 2010, Roger Clemens was indicted by a federal grand jury. Barry Bonds is scheduled for trial in April 2011 (Schmidt (2010)).

III. Career Highlights

America's past-time has always had controversies, whether be it the "Black Sox" cheating scandal of the 1919 World Series, Pete Rose's betting on baseball, or the current allegations of steroid enhanced record breaking power numbers. No player's descent into the steroid controversy may be more polarizing or tragic than Barry Bonds, because no player may have ever needed enhancement less. Bond's pre-controversy career is Hall of Fame worthy. As

background to better understand the implications of our empirical findings, we provide the following brief chronology of Bonds' baseball career and steroid allegations.

a) Career Highlights

Barry Bonds was born on July 24, 1964 in Riverside, California. He debuted in the Pittsburgh Pirates' organization in the spring of 1986. While with the Pirates, Bonds was twice named All-Star, twice named the National League (NL) Most Valuable Player (MVP), was MLB Player of the Year, received three Gold Gloves for his defensive performance in the outfield, and three Silver Slugger awards for his accomplishments at the plate. In 1990, 1991 and 1992 he led the Majors in OPS. While with the Pirates, Bonds went on the disabled list (DL) only once, missing games in June of 1992. He was an everyday player known for his combination of defensive and offensive prowess. After the 1992 season, Bonds ended his career in Pittsburgh and signed the largest contract in MLB history (to that point) with the San Francisco Giants.

Bonds debuted with the San Francisco Giants in Spring 1993. From 1993 to 1998, he amassed six All Star selections, four Gold Gloves, three Silver Sluggers and one NL MVP. In 1993, he became the second player (after Jose Canseco) to have 40 stolen bases and 40 home runs in the same season. In April 1999 at the age of 34, Bonds experienced his first significant injury and had his second visit to the DL in his 13 year career due to injuring the connective tissue in his elbow. This injury, which sidelined him for parts of April and June and all of May, fueled speculation that age had finally caught up with him. After missing his first All-Star game in seven years, Bonds came back in 2000 ready to play again.

In the summer of 2000 Bonds turned 36 and hit the most home runs he had ever hit, 49. At an age when most players' power numbers are significantly declining, Bonds' numbers were climbing. In 2001, Bonds' power numbers went from great to astronomical as he shattered the

three year old record of Mark McGwire for most home runs (73) in a single season. His OPS was over 1.35. From 2001 to 2004, Bonds won four consecutive NL MVP awards and four more Silver Sluggers. However, his days of Gold Gloves were over as his abilities in left field had declined. Bonds' home runs continued to come at a blistering pace, hitting 46, 45 and 45 in 2002, 2003 and 2004, respectively. Bonds' OPS continued to stay high at 1.381, 1.278 and 1.422 in each of these years. While most players of his age experienced a decline in output, Bonds was shattering convention by getting better as he got closer to 40.

In 2005, Bonds missed all but 14 games with a knee injury and his performance began to decline. As rumors of steroid use intensified, many speculated that his injury and decline were related to years of PED usage that had extended his career and inflated his numbers. While Bonds has never been confirmed to have failed a test, his unusually high production after age 35 led to constant speculation. In 2006, when Bonds returned to baseball, he had only one record left in his sights, the all-time home run record of Hank Aaron. In his last three years with the Giants he hit 59 total home runs. During this time Bonds surpassed Hank Aaron's home run mark to finish his career with a record 762 home runs.

b) Steroid Allegations

In 1998, on the heels of the Mark McGwire and Sammy Sosa home run chase, Bonds reconnected with his childhood friend Greg Anderson. Bonds allegedly was upset that while he was playing great baseball, his accomplishments were being overshadowed by power hitters. It is speculated that the home run chase of 1998 and the reconnection with Greg Anderson are the reason that Bonds began to use steroids (Fainaru-Wada and Williams, 2006).

While according to many sources, steroid use in baseball was rampant even before 1998, Bonds remained out of the headlines until 2003. In June 2003, Victor Conte, owner of the Bay

Area Laboratory Co-Operative (BALCO), encouraged Bonds to describe his fitness regimen to *Muscle and Fitness Magazine* (Smaltz, 2003). In the article, Bonds described using Conte's nutrient measuring technique and taking supplements provided by BALCO. Bonds said that this process began in 2000. However, when questioned about Conte and BALCO in 2004 Bonds responded "I don't know BALCO, dude" (ESPN, 2007). In September of 2003, investigators raided BALCO's labs and seized financial and medical records from Conte's company. Conte was under investigation for tax fraud at the time. Also in September, investigators raided the home of Bonds' trainer Greg Anderson and seized documents that supposedly show Bonds was using PEDs.

Toward the end of 2003 a Federal investigation into BALCO began. In December 2003, Bonds twice appeared before a grand jury to testify about BALCO. In the first interview he admitted to using a cream and clear product from Greg Anderson, who was working with BALCO, but was led to believe that they were both permissible products such as flaxseed oil. Later in the month Bonds testified about his relationship with Greg Anderson, admitting paying Anderson for weight training and giving him a bonus after his record setting 73 home runs in 2001. Bonds, however, denied that the documents seized from Anderson's house, such as a calendar containing his name and notes about PEDs, could exist. Later during the proceedings, Gary Sheffield, another MLB player, admitted to receiving the "cream and clear" products from Anderson as well as some pills called red beans. In February 2004, Victor Conte, Greg Anderson and two other men were indicted on charges of money laundering and distributing illegal drugs to athletes. In March of 2004, the San Francisco Times reported that Jason Giambi, Gary Sheffield and others received steroids from BALCO. However, two lawyers for Greg Anderson say Bonds was offered steroids and never took them. Later in April 2004, Conte's

lawyers acknowledged a government memorandum stating that Conte provided steroids to athletes, including Giambi, Sheffield and Bonds, but denied that Conte actually admitted to providing the steroids. Later in 2004 both Sheffield and Giambi admitted that they used steroids between 2001 and 2004 and that those steroids were obtained through Greg Anderson (Fainaru-Wada and Williams, 2006; ESPN, 2007).

In 2005, a former girlfriend of Bonds purported that he told her his 1999 elbow injury was steroid related. Due to this testimony, speculation began to surface that Bonds might be indicted for perjury regarding his grand jury testimony. In 2005, while rehabbing a knee injury, Bonds worked out again with Greg Anderson who was awaiting trial for steroid distribution. In July 2005, Conte and Anderson both pled guilty to conspiracy to distribute steroids (ESPN, 2007).

In 2006, Fainuru-Wada and Williams published the *Game of Shadows: Barry Bonds, BALCO, and the Steroids Scandal that Rocked Professional Sports,* the book which details most fully the case against Bonds. Fainura-Wada and Williams (2006) allege that Bonds began using some steroidal supplements as early as 1997, but that after the McGwire-Sosa home run race of 1998 he began using steroids full time. They say that his regimen included two designer steroids referred to as "the cream" and "the clear," as well as insulin, a human growth hormone which MLB did not test for, testosterone decanoate (a fast-acting steroid known as “Mexican beans”), and trenbolone, a steroid created to improve the muscle quality of cattle. In 2006, MLB launched a special investigation into PEDs in baseball. Finally, in 2007, a grand jury charged Bonds with four counts of perjury and one count of obstruction of justice for his 2003 testimony in which he stated that he had never knowingly taken steroids (ESPN, 2007).

Schmotzer, Switchenko, and Kilgo (2008) use the 2003-2004 seasons as possible steroid usage years in their tests for Bonds, but they also note that steroid usage may have started as early as 2000. They based these assumptions on the Mitchell Report. Based on the available information it is difficult to hypothesize when (alleged) steroid usage might have caused a structural break. It could have been around 2000 or maybe around 2003 to 2004.

IV. The Data

Monthly data on Bonds' hitting production in terms of OPS were collected from BaseballReference.com beginning in 1986, his first year in the majors, through his final year in 2007. All months (September 1994 (MLB strike); May 1999, and April-August 2005 (on the disabled list for all of these months)) in which Bonds did not have any at bats, and the first month of his career (April 1986), in which he had only one at bat, were excluded. Excluding these months, Bonds' mean monthly number of games played was 24.1 and the mean monthly number of at bats was 79.4. Given that his career spanned 132 months and we excluded 8 months due to injury or little playing time, we have 124 months of data with which to perform our time series analysis.

In our analysis, we will examine two aspects of Barry Bonds' hitting performance: actual OPS and the OPS residuals constructed from the Fair (2008) model. The Fair (2008) model has an estimated peak age of 27.59 years (δ). Fair (2008) estimates γ_1= -0.001618 and γ_2= -0.000508 in his Table 1. We calculate β_1= 0.0892812 (=-2*(-0.001618)*27.59) and β_2=0.0280314 (=-2*(-0.000508)*27.59) using Equation 2. These estimated coefficients are the same for all players. Barry Bonds' unique constants (α_1 and α_2) are calculated from the three constraints in equation (2). Bonds' α_2 = 0.699 is estimated by Fair (2008) and α_1 can be calculated by solving equation

(2), which yields $\alpha_1 = -0.14595$ (=0.699 + (-0.000508-(-0.001618))*27.59^2). Bonds' predicted OPS before age 27.6 is calculated by the first equation in (1) using the above values for the coefficients. Using the second equation in (1), Bonds' OPS after age 27.6 can be calculated as well. The OPS residuals are calculated as the difference between the actual OPS and the predicted OPS. The data on actual OPS and the OPS residuals are displayed graphically in Figure 1. From Figure 1, it is apparent that the actual values of OPS during the later part of Bonds' career are much higher than the predicted OPS numbers. To more accurately assess the properties of Bonds' OPS data, in the next section we will perform times series tests with structural breaks.

V. Time Series Tests with Structural Breaks

Since a baseball player's OPS generally rises, reaches a peak, and then declines over his career, it should not be expected that an OPS time series would be stationary in terms of having a constant mean or constant trend. In contrast, we would expect the actual OPS series to have a break near the peak of the productivity profile (around age 27 or 28 for most players). Then, following this peak, we would expect the productivity to decline gradually. In the case of Fair's (2008) OPS residuals, we would expect these to be randomly distributed around zero. In particular, we should not expect the OPS residuals to have any trend or structural break. We will test these predictions by conducting time series tests on Bonds' actual OPS and OPS residuals to determine if these series are stationary and determine if and when a structural break occurs. To do so, we utilize the endogenous break LM unit root test of Lee and Strazicich (2003). This test can be utilized to identify two breaks in level and trend while jointly testing the null hypothesis of a unit root. A noted feature of the endogenous LM unit root tests is that they are not subject to

spurious rejections and will be valid whether a break occurs under the null or alternative hypothesis (Lee and Strazicich, 2001, 2003).[1]

Our testing methodology can be briefly described as follows. Based on the LM (score) principle, a unit root test statistic can be obtained from the following regression:

$$\Delta y_t = \delta'\Delta Z_t + \varphi\tilde{S}_{t-1} + \Sigma\gamma_i\Delta\tilde{S}_{t-i} + \varepsilon_t \,, \qquad (3)$$

where $\tilde{S}_t$ is a de-trended series such that $\tilde{S}_t = y_t - \tilde{\psi}_x - Z_t\tilde{\delta}$, t = 2,..,T, $\tilde{\delta}$ is a vector of coefficients in the regression of Δy_t on ΔZ_t, $\tilde{\psi}_x = y_1 - Z_1\tilde{\delta}$, where y_1 and Z_1 are the first observations of y_t and Z_t, respectively, Δ is the difference operator, and ε_t is an *iid* error term with zero mean and finite variance. The first differenced lagged terms, $\Delta\tilde{S}_{t-i}$, i = 1,..,k, are included as necessary to correct for serial correlation. In the model with two level and trend breaks, Z_t is described by [1, t, D_{1t}, D_{2t}, DT_{1t}^*, DT_{2t}^*]′, where $D_{jt} = 1$ for $t \geq T_{Bj} + 1$, j = 1, 2, and zero otherwise; $DT_{jt}^* = t - T_{Bj}$ for $t \geq T_{Bj} + 1$, j = 1, 2, and zero otherwise; and T_{Bj} is the time period of the breaks.[2] Under the unit root null hypothesis, $\varphi = 0$ in (3) and the test statistic is denoted as:

$$\tilde{\tau} = \text{t-statistic for the null hypothesis } \varphi = 0. \qquad (4)$$

Rejection of the null implies a "trend-break stationary" time series ($\varphi < 0$), where the series is stationary around one or two changes in the level and trend slope.

We adopt a general-to-specific procedure to determine the number of lagged first differenced terms ($\Delta\tilde{S}_{t-i}$) that correct for serial correlation in (3). At each combination of break points, $\lambda = (\lambda_1, \lambda_2)$ ′, where $\lambda_j = T_{Bj}/T$, j=1, 2, in the time interval [.1T, .9T] to eliminate end points, we begin with a maximum number of k = 8 lagged first differenced terms and examine the last term to see if it is significantly different from zero at the 10% level (critical value in an

asymptotic normal distribution is 1.645). If not significant, the maximum lagged term is dropped and the model re-estimated with k = 7 terms, etc. The procedure is repeated until either the maximum lagged term is found or k = 0, at which point the procedure stops. This type of procedure has been shown to perform well as compared to other data-dependent procedures to select the number of lagged first differenced terms (e.g., Ng and Perron, 1995).[3]

The test results are displayed in Table 1. The null of a unit root is rejected at the 1% level of significance in each series. In each series, the identified structural breaks occur in June 1993 and September 2000, which corresponds to an age of 28.9 years and 36.1 years, respectively. Given that the actual OPS and residual OPS are each stationary around two breaks, we estimated a simple OLS regression on three intercepts and trends before and after June 1993 and September 2000, respectively. We denote these intercept (D_t) and trend (T_t) variables as $D_{1986:05\text{-}1993:06}$, $D_{1993:07\text{-}2000:09}$, $D_{2000:10\text{-}2007:09}$, and $T_{1986:05\text{-}1993:06}$, $T_{1993:07\text{-}2000:09}$, and $T_{2000:10\text{-}2007:09}$, respectively. The regression results are displayed in Table 2.

To visually examine our findings, we constructed plots of the actual OPS and residual OPS series along with their fitted values from the regressions in Table 2. The plots are displayed in Figures 2 and 3. In Figure 2, we observe a positive trend from the age of 21.8 (May 1986) to 28.9 (June 1993). As noted, Fair (2008) found that the OPS is expected to peak at age 27.6, which corresponds roughly with Bonds' age of 28.9 in June 1993. Then, Bonds' OPS follows a stable mean with an approximate zero trend slope until the age of 36.1 in September 2000, at which point his OPS dramatically increases and then declines thereafter until his retirement in September 2007. Note that we observe evidence of a plateau in the actual OPS after age 28.9, where the slope is not statistically different from zero in the interval from age 28.9 to age 36.1. This outcome is somewhat expected, although slightly different from Fair's (2008) prediction of

a slight decline beginning around this age for MLB players. In Figure 3, Bonds' residual OPS series also displays a plateau from the age of 28.9 to age 36.1, hovering around zero as would typically be expected.

Most dramatic is the relatively large increase in Bonds' actual OPS starting at the age of 36.1 in September 2000. His productivity remains high, while declining thereafter until his retirement at age 43. A similar upward shift is observed in Bonds' OPS residuals. Only after this large upward shift does the residual series exhibit a gradual decline. These upward trend shifts are unexpected for a hitter in the 36 to 43 year age range.

VI. Conclusion

We examine time series on Barry Bonds' OPS by utilizing unit root tests that endogenously determine one or two structural breaks in level and trend. Our test results indicate that Bonds' actual OPS and residual OPS are trend-stationary with two breaks in June 1993 and September 2000. Based on the timing of the breaks, there are a few notable conclusions. First, in June 1993, we find evidence of a break in the actual OPS at the peak age of 28.9 after which his OPS plateaus, similar to what one would expect from Fair's (2008) model. Second, the OPS residuals are not scattered around zero, as would be expected if Bonds' performance followed the typical pattern of 10 year plus career players in MLB. Instead, Bonds' actual OPS and OPS residuals increased dramatically at the age of 36. Third, it is interesting to observe that the timing of this second break, September 2000, coincides closely with the time when Bonds' stated he began working with Greg Anderson and BALCO. This is also close to the time of possible steroid use mentioned in the Mitchell report. Moreover, 2001 is the time period when Bonds broke Mark McGwire's home run record and won the first of four consecutive NL MVP awards.

If Bonds began using steroids sometime around the 1999 season, as reported by Fainaru-Wada and Williams (2006), it is plausible, especially considering Bonds' injury during the 1999 season that the full effects of the new training regimen and PEDs might not have come to fruition until late in the 2000 season. While there is certainly no conclusive evidence that Bonds used PEDs, given the timing of the break, one can only speculate if the 'training and products' provided by Anderson and BALCO were the cause.

The intrigue surrounding steroid use continues to grow. On January 11, 2010, in a statement released via the St. Louis Cardinals' website, Mark McGwire admitted to using PEDs beginning in 1989 or 1990. By his own admission he used steroids on and off from the late 1980s through the 1990s. In his statement McGwire claims that his use of steroids was mostly for recovery from injuries. This admission, which was not surprising in its content, lends credence to the idea that Bonds could have been influenced by this suspicion, perhaps by knowing what other players, like Jose Canseco, claim to have known at the time. As more suspected steroid users come forward, perhaps some clarity will be shed on the reality of Bonds' performance during the late 1990s and early 2000s.

References

Addona, Vittorio and Jeremy Roth (2010). "Quantifying the Effect of Performance-Enhancing Drug Use on Fastball Velocity in Major League Baseball," *Journal of Quantitative Analysis in Sports,* Volume 6, Issue 6, Article 6, 1-14.

DeVany, Art (2010), "Steroids and Baseball." *Economic Inquiry*, forthcoming.

Dinardo, John and Jason Winfree (2010). "The Law of Genius and Home Runs Refuted." *Economic Inquiry*, 48, No. 1, 51–64.

Eber, Nicolas (2008). "The Performance-Enhancing Drug Game Reconsidered: A Fair Play Approach.**"** *Journal of Sports Economics, 9*: 318-327.

ESPN (2007). "Barry Bonds Steroids Timeline." http://sports.espn.go.com/mlb/news/story?id=3113127, December 7.

Fainaru-Wada, Mark, and Lance Williams (2006). *Game of Shadows: Barry Bonds, BALCO, and the Steroids Scandal that Rocked Professional Sports*. New York, NY: Gotham Publishing.

Fair, Ray C. (1994). "How Fast Do Old Men Slow Down?" *The Review of Economics and Statistics*, *76*, 103–118.

Fair, Ray C. (2007). "Estimated Age Effects in Athletic Events and Chess." *Experimental Aging Research, 33*, 37–57.

Fair, Ray C. (2008). "Estimated Age Effects in Baseball." *Journal of Quantitative Analysis in Sports, 4 (1)*, Article 1, 1-39.

Fort, Rodney and Young Hoon Lee (2006). "Stationarity and Major League Baseball Attendance Analysis." *Journal of Sports Economics*, *7 (4)*, 408–415.

Fort, Rodney and Young Hoon Lee (2007). "Structural Change, Competitive Balance, and the Rest of the Major Leagues." *Economic Inquiry*, 45 (3), 519-532.

Harrington, Joseph E., Jr. (2009). *Games, Strategies, and Decision Making*. New York: Worth Publishers.

Haugen, K. (2004). "The Performance-Enhancing Drug Game." *Journal of Sports Economics, 5*, 67-86.

Lee, Young Hoon, and Rodney Fort (2005). "Structural Change in Baseball's Competitive Balance: The Great Depression, Team Location, and Racial Integration." *Economic Inquiry*, 43, 158-169.

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.

Leeds, Michael A, and Peter von Allmen (2007). *The Economics of Sports*. Third Edition. Boston: Pearson Education Inc.

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.

Scully, G. W. (1989). *The Business of Baseball*. Chicago: University of Chicago Press.

Scully, G. W. (1995). *The Market Structure of Sports*. Chicago: University of Chicago Press.

Schmidt, M. B., and D. J. Berri (2003). "On the Evolution of Competitive Balance: The Impact of an Increasing Global Search." *Economic Inquiry*, 41, 692-704.

Schmidt, M. B., and D. J. Berri (2004). "The Impact of Labor Strikes on Consumer Demand: An Application to Professional Sports." *American Economic Review*, 94, 344-357.

Schmidt, Michael S. (2010). "Clemens Lied about Doping, Indictment Charges." New York Times. August 20, A1.

Smaltz, Jim (2003). "The King of Swing." *Muscle and Fitness*, 64(6), 170-193.

Schmotzer, Brian J., Jeff Switchenko, and Patrick D. Kilgo (2008). "Did Steroid Use Enhance the Performance of the Mitchell Batters? The Effect of Alleged Performance Enhancing Drug Use on Offensive Performance from 1995 to 2007," *Journal of Quantitative Analysis in Sports* Volume 4, Issue 3, Article 4: 1-14.

Sommers, P. (2008). "The Changing Hitting Performance Profile in Major League Baseball, 1966-2006." *Journal of Sports Economics, 9,* 435-440.

TABLE 1. LM Unit Root Test Results of Barry Bonds Actual OPS and Residual OPS, May 1986 – September 2007 (124 monthly observations)

Country	*k*	$\hat{T}_B$	*Test Statistic*	*Break Points*
Actual OPS	0	June 1993; September 2000	-11.8254***	λ = (.4, .8)
Residual OPS	0	June 1993; September 2000	-11.8764***	λ = (.4, .8)

Notes: The data omits the first month and the seven months where Barry Bonds did not play. 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. $\hat{T}_B$ denotes the time periods of the identified breaks. The critical values are shown below and come from Lee and Strazicich (2003). 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.

Break points	Critical values		
$\lambda = (T_{B1}/T, T_{B2}/T)$	1%	5%	10%
$\lambda = (.2, .4)$	–6.16	-5.59	-5.27
$\lambda = (.2, .6)$	–6.41	-5.74	–5.32
$\lambda = (.2, .8)$	–6.33	-5.71	–5.33
$\lambda = (.4, .6)$	–6.45	-5.67	–5.31
$\lambda = (.4, .8)$	-6.42	-5.65	–5.32
$\lambda = (.6, .8)$	-6.32	-5.73	–5.32

Table 2. OLS Regression Results of Barry Bonds' Actual OPS and Residual OPS on Two Level and Trend Breaks, May 1986 – September 2007

$$\text{Actual OPS}_t = 0.6937D_{1986:05\text{-}1993:06} + 1.0209D_{1993:07\text{-}2000:09} + 1.4802D_{2000:10\text{ -}2007:09}$$

(16.268) (16.869) (20.618)

$$+ 0.0090T_{1986:05\text{-}1993:06} + 0.0012T_{1993:07\text{-}2000:09} - 0.0124T_{2000:10\text{ -}2007:09} + e_t$$

(4.521) (0.497) (-3.854)

Adjusted R-squared = 0.434 SER = 0.192 Q(12) = 6.521 (p-value = 0.888)

$$\text{Residual OPS}_t = - 0.3477D_{1986:05\text{-}1993:06} - 0.0689D_{1993:07\text{-}2000:09} + 0.4260D_{2000:10\text{ -}2007:09}$$

(-8.148) (-1.136) (5.993)

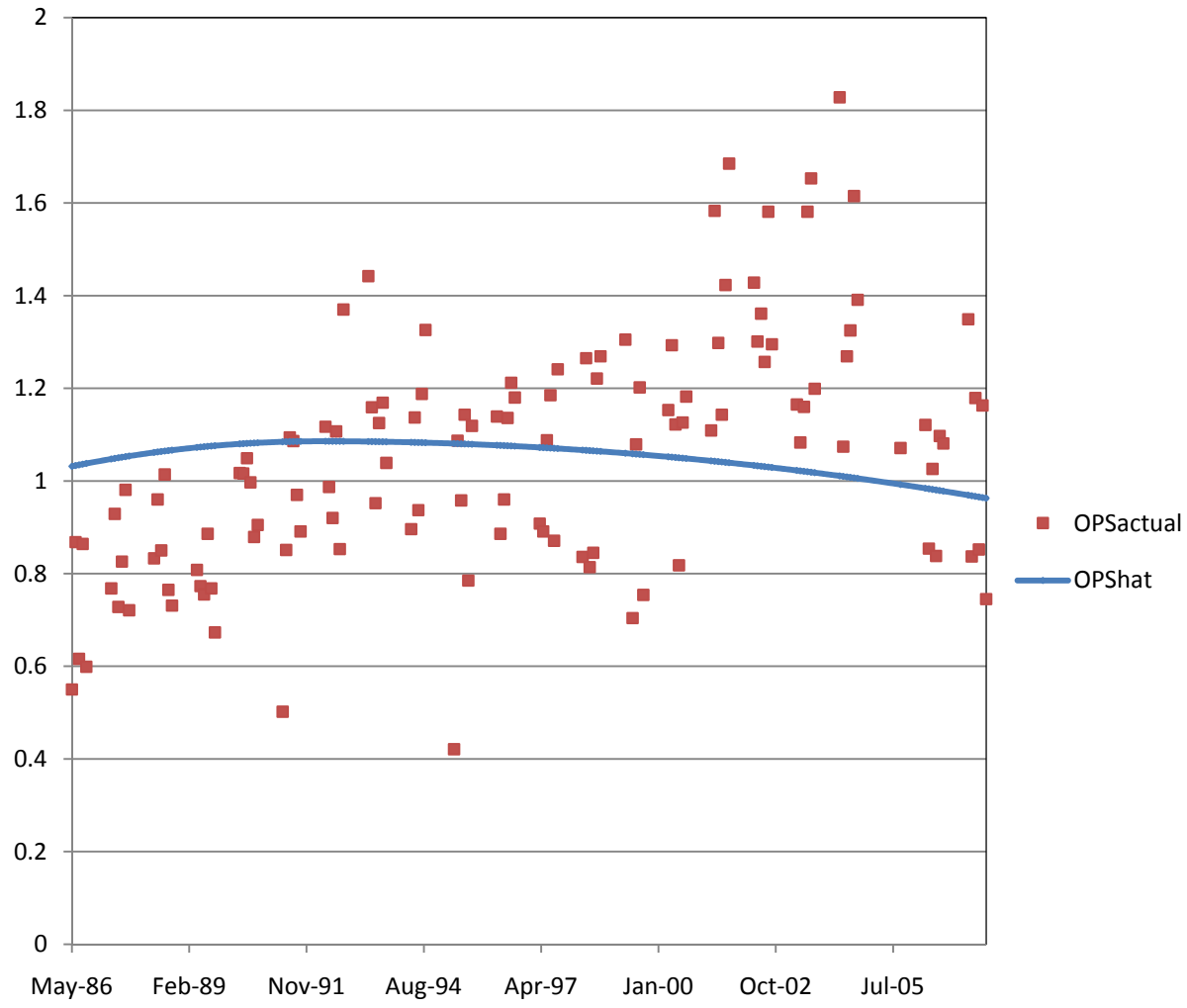
$$+ 0.0077T_{1986:05\text{-}1993:06} + 0.0021T_{1993:07\text{-}2000:09} - 0.0100T_{2000:10\text{ -}2007:09} + e_t$$

(3.860) (0.873) (-3.122)

Adjusted R-squared = 0.476 SER = 0.192 Q(12) = 6.517 (p-value = 0.888)

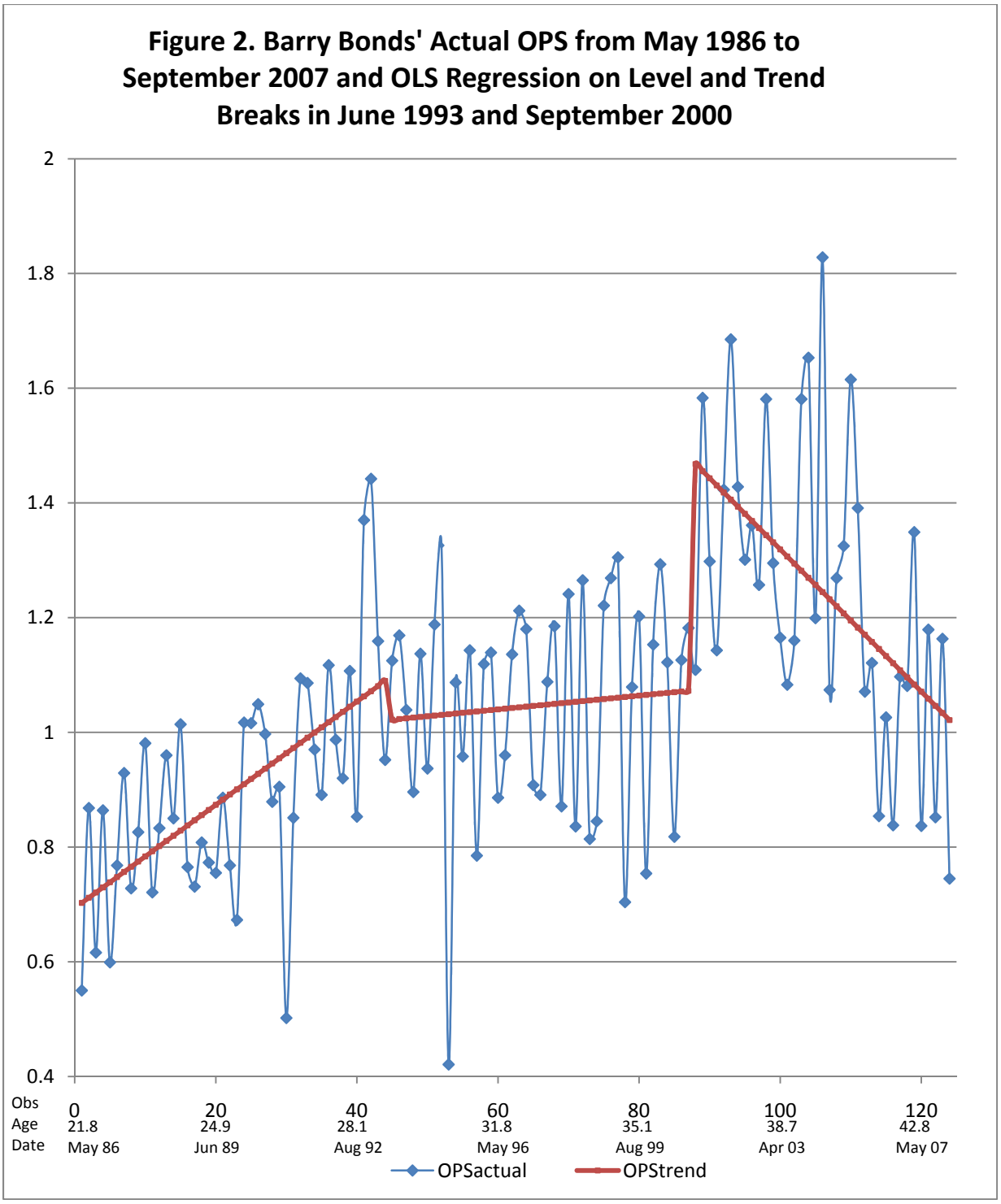
Notes: t-statistics are shown in parentheses. White's robust standard errors were utilized to control for heteroskedasticity. The Ljung-Box Q-statistic for 12 lags is denoted by Q(12) and indicates that the null of no serial correlation cannot be rejected at the usual significance levels.

Figure 1: Bonds' Actual OPS vs. Predicted OPS, May 1986-September 2007



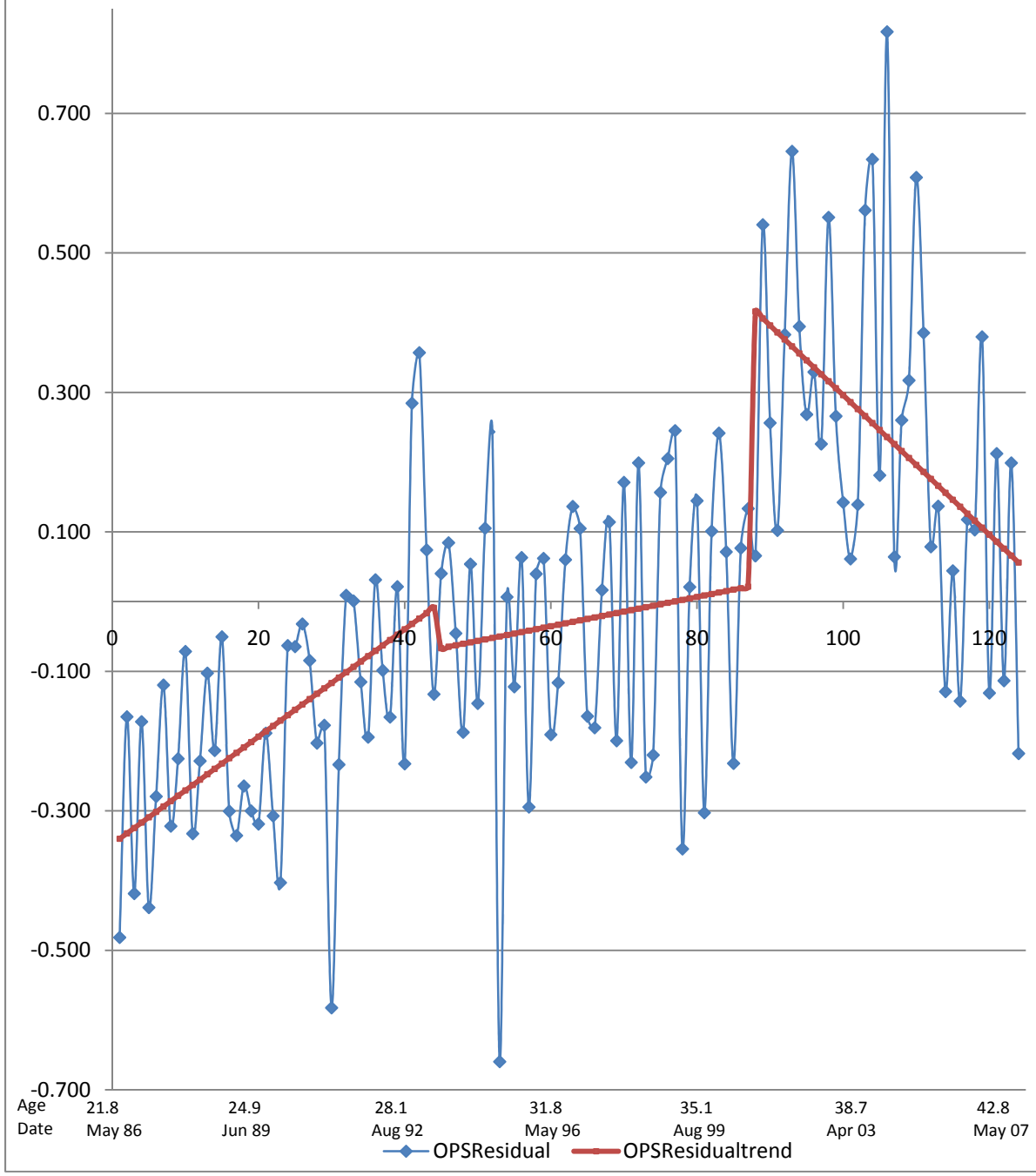
Notes: OPShat denotes the OPS values predicted from equations (1) and (2).

Figure 2. Barry Bonds' Actual OPS from May 1986 to September 2007 and OLS Regression on Level and Trend Breaks in June 1993 and September 2000



Notes: The data omit the first month and the seven months when Barry Bonds did not play

Figure 3. Barry Bonds' Residual OPS from May 1986 to September 2007 and OLS Regression on Level and Trend Breaks in June 1993 and September 2000



Notes: The data omit the first month and the seven months when Barry Bonds did not play.

Endnotes

[1] Other than the work of Scully (1995), Fort and Lee (2006) note that time series tests have only recently become more widely applied to sports data (e.g., Schmidt and Berri, 2003, 2004; Lee and Fort, 2005; Fort and Lee, 2007).

[2] Note that the test regression in (3) includes ΔZ_t instead of Z_t, where ΔZ_t is $[1, B_{1t}, B_{2t}, D_{1t}, D_{2t}]'$ with $B_{jt} = \Delta D_{jt}$ and $D_{jt} = \Delta DT_{jt}^*$, $j = 1, 2$.

[3] The Gauss code for the two-break minimum LM unit root test is available on the web site: http://www.cba.ua.edu/~jlee/gauss.