

Stationarity and Major League Baseball Attendance Analysis

RODNEY FORT

Washington State University

YOUNG HOON LEE

Hansung University

If a sports time series, such as attendance, is nonstationary, then the use of level data (e.g., demand estimation using panel data) leads to biased estimates, and the direction of the bias is unknown. In past works, authors have failed to reject nonstationary data, taken first differences, and proceeded with further analysis. That is a legitimate approach, although limiting (e.g., no elasticity estimates can be had from first differences). However, if the data are stationary, then all is well with the usual applications to level data (e.g., taking logs gives direct elasticity estimates). This article rejects that the Major League Baseball attendance time series is nonstationary with break points and suggests the break points deserve additional analysis to facilitate attendance demand investigations.

Keywords: baseball attendance; sports time series; stationarity

It wasn't until relatively recently that Scully (1995) formally applied time series techniques to sports data in his work on winning cycles. And the work assessing the time series behavior of sports data related to attendance and competitive balance all has occurred within the past 10 years (Davies, Downward, & Jackson, 1995; Dobson & Goddard, 1998; Y. H. Lee & Fort, 2005; Schmidt, 2001; Schmidt & Berri, 2001, 2002, 2003, 2004; Simmons, 1996).

In approaching sports time series such as measures of competitive balance or attendance, the important first question is whether the time series is stationary. The answer to that question has implications for additional analysis of the time series, itself, and for practitioners wishing to use the time series for regression purposes, for example, annual demand regressions. If a sports time series, such as attendance,

is nonstationary, then the use of level data (e.g., demand estimation using panel data) leads to biased estimates and the direction of the bias is unknown. A technique such as taking first differences is a useful approach to a nonstationary series, although limiting (e.g., no elasticity estimates can be had from first differences). However, if the data are stationary, then all is well with the usual applications to level data (e.g., taking logs gives direct elasticity estimates).

It ends up that the past literature on attendance, using ordinary unit-root tests, has failed to reject nonstationary time series and taken the first differences approach. We reverse that finding using unit-root tests with break points for attendance in Major League Baseball (MLB). We recognize that our contribution in this note is only a partial one concerning break point analysis. In this note, we do not proceed ahead to identify more precisely when break points occur and the sign and significance of their impacts. Nor do we take break points into account and offer any panel data analysis of attendance demand. The note itself is long enough, and it seemed useful to bring the result to the attention of those analysts interested in time series analysis and panel-data attendance estimation. In addition, we offer a schematic approach to most of these issues.

The note proceeds as follows. The following section demonstrates the issues schematically and in more detail. In the third section, we discuss the two steps to unit-root testing with break points for MLB attendance. The following section details the results of our findings for the first step, reversing past findings. Conclusions round out the note in the last section.

SCHEMATIC DEMONSTRATION OF THE ISSUES

Figure 1 is a simple schematic of a general approach to the investigation of the nonstationary behavior of sports time series. In approaching a given time series, the first best practice is an ordinary unit-root test of nonstationary behavior, labeled Step 1 in Figure 1. If the test rejects nonstationary behavior—Step 1.1 along Path 1 in Figure 1—the methods developed by Bai and Perron (1998, 2003; henceforth, the BP method) are valid. An example of this type of result down Path 1 is in the competitive balance analysis of Y. H. Lee and Fort (2005). In addition, along this path, level-data practitioners may proceed to use the entire time series, down Step 1.2. For example, it is valid to estimate attendance regressions on the entire sample.

However, if the ordinary unit-root tests fail to reject nonstationary behavior, suggesting Path 2 in Figure 1, further insight can be gained about the behavior of the time series using unit-root tests with endogenous structural breaks, Step 2.1. This technique is detailed in the third section and applied to MLB attendance data in the fourth section. If this type of subsequent unit-root test with break points rejects nonstationary behavior, the BP method once again comes into play, Step 2.1A. Unlike Path 1, however, the next Step 2.1B allows level-data analysis only where the data are stationary, that is, between specific break points in the time series. Finally, if the unit-root test with break points fails to reject a nonstationary time

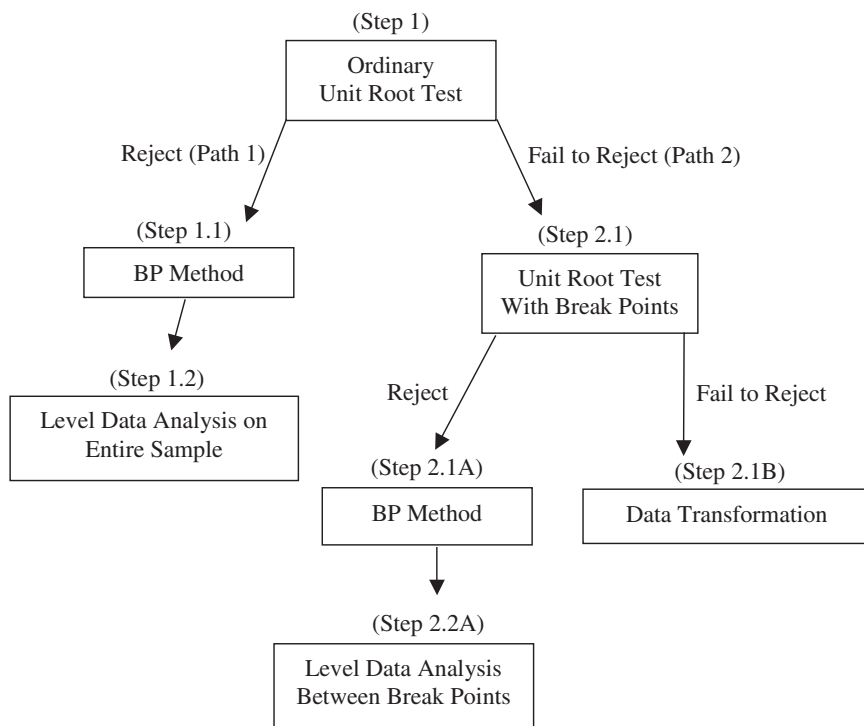


Figure 1: Time Series Approach Schematic

series with break points, Step 2.1B shows a data transformation such as taking first differences before additional time series analysis.

We stress the two-step nature of break point analysis because it defines the limits of what we attempt in this note. Although we formally describe both steps in the next section, we only do Step 1 and Step 2.1 in this note. For Step 2.1: Is there evidence that break points matter, so that unit-root tests with break points indicate Step 2.1A? However, after Step 2.1, all we know is whether the data are stationary with break points. If the data are stationary with break points, the BP method in Step 2.1A still needs to be employed to precisely determine the parameters of the break points—with statistical precision, when do the break points occur and, subject to the limitations of the methodology, what is the direction of their impact on attendance over time? Our previous experience in Y. H. Lee and Fort (2005) makes it clear that actually doing Step 2.1A and then 2.2A will be a lengthy paper in and of itself. Knowing what we can from Step 2.1 is the point of this note.

We comment finally on the relationship of this so-called two-step approach to past work on the analysis of time series. At Step 1, Schmidt and Berri (2001) found

the American League (AL) and National League (NL) annual attendance series are nonstationary (integrated of order one). They then move immediately to Step 2.1B, bypassing Step 2.1 altogether. This is a legitimate econometric choice; however, we point out that doing Step 2.1, the unit-root test with break points, may yield additional insight into the time series behavior of the data. And if the time series is stationary with break points, proceeding to Step 2.1A may alleviate the need to take first differences prior to analyzing annual attendance data.

UNIT-ROOT TESTS OF ENDOGENOUS BREAK POINTS IN MLB ATTENDANCE

In this section, we go into more detail on Step 2.1 only. At Step 1, in pioneering work, Perron (1989) show that standard unit-root tests lose power when the stationary alternative is true and an existing structural break is ignored. In the earliest development of Step 2.1, Perron (1989) uses a modified Dickey-Fuller unit-root test that includes dummy variables to account for one known, or exogenous, structural break. The break point of the trend function is fixed (exogenous) and chosen independently of the data. Nelson and Plosser (1982) found evidence in favor of the unit-root hypothesis for 13 of 14 long-term annual macro series. Perron (1989) reverses the Nelson and Plosser conclusions in 10 of the 13 series.

Subsequent literature, including Banerjee, Lumsdaine, and Stock (1992), J. Lee and Strazicich (2001), Perron (1997), Vogelsang and Perron (1998), and Zivot and Andrews (1992, henceforth ZA), incorporates an endogenous break point. This endogenous break point literature has focused on testing the unit-root null hypothesis against a one-break alternative. However, there may be more than one break point for longer time series, and Lumsdaine and Papell (1997, henceforth LP) extend the endogenous break methodology to allow for a two-break alternative. LP found that results regarding tests of the unit-root hypothesis are sensitive to the number of breaks in the alternative specification even within the class of endogenous break models.

Unlike Perron's (1989) null hypothesis, ZA (1992) and LP (1997) assume no breaks under the unit-root null and derive their critical values accordingly. Thus, rejection of the null does not necessarily imply rejection of unit root *per se* but would imply rejection of a unit root without breaks. Nunes, Newbold, and Kuan (1997) show that this assumption leads to size distortions in the presence of a unit root with break. J. Lee and Strazicich (2001, 2003) further investigate this issue and discover the source of the size distortions. In the presence of a unit root with a structural break, the ZA and LP tests tend to select the break point where bias and size distortions are the greatest. J. Lee and Strazicich (2003, henceforth LS) propose a two-break minimum Lagrange multiplier (LM) unit-root test in which the alternative hypothesis unambiguously implies the series is trend stationary.

Perron (1989) considers three structural break models. The crash model allows for a one-time change in level. The changing-growth model allows for a change in

trend slope. The final model allows for a change in the level and trend. In what follows, we employ this last, most general specification.

The two-break minimum LM unit-root test can be described as follows. According to the LM principle, a unit-root test statistic can be obtained from the following regression:

$$\Delta y_t = d' \Delta Z_t + \phi \tilde{S}_{t-1} + \Sigma \gamma_t \Delta \tilde{S}_{t-1} + \varepsilon_t. \quad (1)$$

Δ is the difference operator and \tilde{S}_t is a detrended series such that $\tilde{S}_t = y_t - \tilde{\psi}_x - Z_t \tilde{\delta}$, $t = 2, \dots, T$. $\tilde{\delta}$ is a vector of coefficients in the regression of Δy_t on ΔZ_t and $\tilde{\psi}_x = y_1 - Z_1 \tilde{\delta}$. ε_t is the contemporaneous error term and is assumed i.i.d. $N(0, \sigma^2)$. Z_t is a vector of exogenous variables. Corresponding to the two-break equivalent of Perron's (1989) most general model (levels and trends allowed to vary), with two changes in level and trend, Z_t is described by $[1, t, D_{1p}, D_{2p}, DT_{1p}, DT_{2p}]'$, where $D_{jt} = 1$ for $t \geq T_{Bj} + 1$, $j = 1, 2$, and zero otherwise, $DT_{jt} = 1$ for $t \geq T_{Bj} + 1$, $j = 1, 2$, and zero otherwise, and T_{Bj} stands for the time period of the breaks.

The unit-root null hypothesis is described in Equation (1) by $\phi = 0$ and the test statistic is a t statistic for this null. To endogenously determine the location of two breaks ($\lambda_j = T_{Bj}/T$, $j = 1, 2$), LS use a grid search to find a minimum t statistic. Therefore, the critical values correspond to the location of the breaks (see LS for more detail and the critical value tables).

To implement this test, LS first determined the number of augmentation terms \tilde{S}_{t-j} , $j = 1, \dots, k$, that correct for serial correlation in Equation (1). At each combination of break points $\lambda = (\lambda_1, \lambda_2)'$ in the time interval $[.1T, .9T]$, where T is the sample size, LS determines k by following a general-to-specific procedure described by Perron (1989). Start with an upper bound k_{\max} for k . If the last included lag is significant, choose $k = k_{\max}$. If not, reduce k by 1 until the last lag becomes significant. If no lags are significant, set $k = 0$.

We repeat the earlier caveats on this overall method, and for the LS version we employ in particular. First, the researcher chooses the number of break points a priori and tests to see if that number of breaks occurs. This means that the number of break points is a rough guess. Second, the point is only to test nonstationary behavior, and subsequent steps may be required to more fully analyze the significance and direction of break points. For level-data practitioners, if the test rejects nonstationary behavior with break points, the BP method could be applied to more precisely determine the break points, their significance, and direction, before moving on to incorporate break points into the analysis of level data (i.e., Step 2.1A in Figure 1).

MLB ATTENDANCE DEMONSTRATION

We demonstrate Steps 1 and 2.1 for MLB attendance, 1900-2003. For Step 1, we conduct Augmented Dickey-Fuller (ADF) and Phillips-Perron (PP) unit-root tests that do not consider structural break points. The results are presented in Table 1.

TABLE 1: Augmented Dickey-Fuller (ADF) and Phillips-Perron (PP) Unit-Root Tests

		American League Attendance	National League Attendance
ADF (<i>p</i>)	Constant	1.350 (2)	0.371 (2)
	Trend	-1.102 (2)	-1.693 (2)
PP (<i>l</i>)	Constant	0.969 (4)	0.288 (4)
	Trend	-2.127 (4)	-1.896 (4)

NOTE: *p*: the number of lags. *l*: lag truncation.

TABLE 2: Two-Break Minimum Lagrange Multiplier (LM) Unit-Root Tests

League	\hat{k}	\hat{T}_B	\hat{t}_{γ_j}	Test Statistic	Critical Value Break Points
American League	0	1965, 1987	-.427, .973	-6.009 ^b	$\lambda = (.65, .85)$
National League	0	1951, 1991	2.178 ^a , 4.559 ^a	-8.216 ^a	$\lambda = (.50, .90)$

NOTE: \hat{k} is the optimal number of lagged first-difference terms included in the unit root test to correct for serial correlation. \hat{T}_B denotes the estimated break points. \hat{t}_{γ_j} is the *t* value of DT_{jt} , for *j* = 1, 2. See J. Lee and Strazicich (2003) Table 2 for critical values.

a. Significant at the 99% critical level.

b. Significant at the 95% critical level.

The number of lags is determined by minimization of the Schwartz-Bayesian criterion for the ADF test and by the truncation suggested by Newey and West (1994) for the PP test. The unit-root hypothesis is not rejected for either the AL or NL. Both of the ADF and PP tests suggest that the league attendance of MLB is nonstationary as found in the literature cited earlier.

Moving on to Step 2.1, the results of the two-break minimum LM unit-root tests are in Table 2 (our thanks to Professor Lee, the *L* in the LS designation in this article, for the GAUSS code). The unit-root null is rejected in the AL and NL attendance series at the 5% and 1% significance levels, respectively. Unlike the NL, the break points in the AL attendance series are not statistically significant. This implies the two-breaks test may not be appropriate for AL because including two breaks instead of one can adversely affect the power to reject the null hypothesis.

We performed additional tests using the one-break minimum LM unit-root test developed in J. Lee and Strazicich (2001). The results are shown in Table 3. The one-break test results for the unit-root hypothesis are essentially the same as the two-break results, except that the unit-root null is rejected in the AL at the stronger 1% significance level rather than at the 5% level, and the trend dummy for a single break is statistically significant in the AL. This implies that the one-break test may be more appropriate for the AL while the two-break test remains more appropriate for the NL.

We also performed ZA (1992) and Perron (1997) tests with one break, and the LP test with two breaks. The results (available on request from the authors) are

TABLE 3: One-Break Minimum Lagrange Multiplier (LM) Unit-Root Test

<i>League</i>	\hat{k}	\hat{T}_B	\hat{t}_{γ_j}	<i>Test Statistic</i>	<i>Critical Value Break Points</i>
American League	0	1975	4.885 ^a	-5.383 ^a	$\lambda = .75$
National League	0	1975	-.353	-6.320 ^a	$\lambda = .60$

a. Significant at the 99% critical level.

unanimously consistent with the results in Tables 2 and 3. Along with the results of the rest of the testing for Step 2.1, this clearly suggests that MLB attendance data are stationary with structural breaks.

The next step, Step 2.1A, in Figure 1 would implement the BP method to actually estimate when the break points occur in time, their sign, and statistically test their magnitudes. We stress again that even though there are break point years shown in Tables 2 and 3, these are only rough estimates, and the BP method is required to bring formal statistical tests to bear on these results. However, this research note has gone far enough, and the BP analysis of attendance is the subject of our other article (Fort & Lee, 2005).

CONCLUSIONS

If the time series data on attendance are nonstationary, the use of level data in panel applications leads to biased estimates, and the direction of the bias is unknown. Authors then have taken first differences and proceeded with further analysis of the time series. That is one econometric approach, although limiting (e.g., no elasticity estimates can be had from first differences). However, if the data are stationary, then all is well with the usual applications to level data (e.g., taking logs gives direct elasticity estimates). We found that the data on MLB attendance are stationary with break points, reversing past works that failed to reject that the data are nonstationary.

The implications are as follows. First, a more precise endogenous determination of the break points, following the BP method, is suggested. Second, practitioners on level data will enhance their results if they use dummy variable techniques that essentially divide the data into stationary subsamples around structural breaks. Otherwise, spurious regression results may be found, and policy making in MLB would be misdirected.

REFERENCES

- 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.

- Banerjee, A., Lumsdaine, R. L., & Stock, J. H. (1992). Recursive and sequential tests of the unit root and trend-break hypothesis: Theory and international evidence. *Journal of Business and Economic Statistics*, 10, 271-287.
- Davies, B., Downward, P., & Jackson, I. (1995). The demand for rugby league: Evidence from causality tests. *Applied Economics*, 27, 1003-1007.
- Dobson, S. M., & Goddard, J. A. (1998). Performance and revenue in professional league football: Evidence from Granger causality tests. *Applied Economics*, 30, 1641-1651.
- Fort, R., & Lee, Y. H. (2005). *Time series behavior of MLB attendance: Structural change and the uncertainty of outcome hypothesis* (Working paper). Pullman: Department of Economics, Washington State University.
- Lee, J., & Strazicich, M. C. (2001). Break point estimation and spurious rejections with endogeneous unit root tests. *Oxford Bulletin of Economics and Statistics*, 63, 535-558.
- Lee, J., & Strazicich, M. C. (2003). Minimum LM unit root test with two structural breaks. *Review of Economics and Statistics*, 85, 1082-1089.
- 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.
- Lumsdaine, R., & Papell, D. (1997). Multiple trend breaks and the unit-root hypothesis. *Review of Economics and Statistics*, 79, 212-218.
- Nelson, C. R., & Plosser, C. I. (1982). Trends and random walks in macroeconomic time series. *Journal of Monetary Economics*, 10, 139-162.
- Newey, W., & West, K. (1994). Automatic lag selection in covariance matrix estimation. *Review of Economic Studies*, 61, 631-653.
- Nunes, L., Newbold, P., & Kuan, C. M. (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.
- Perron, P. (1997). Further evidence on breaking trend functions in macroeconomic variables. *Journal of Econometrics*, 80, 355-385.
- Schmidt, M. B. (2001). Competition in Major League Baseball: The impact of expansion. *Applied Economic Letters*, 8, 21-26.
- Schmidt, M. B., & Berri, D. J. (2001). Competitive balance and attendance: The case of Major League Baseball. *Journal of Sports Economics*, 2, 145-167.
- Schmidt, M. B., & Berri, D. J. (2002). The impact of the 1981 and 1994-1995 strikes on Major League Baseball attendance: A time-series analysis. *Applied Economics*, 34, 471-478.
- Schmidt, M. B., & Berri, D. J. (2003). On the evolution of competitive balance: The impact of an increasing global search. *Economic Inquiry*, 4, 692-704.
- 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.
- Scully, G. W. (1995). *The market structure of sports* Chicago: University of Chicago Press.
- Simmons, R. (1996). The demand for English League Football: A club-level analysis. *Applied Economics*, 28, 139-155.
- Vogelsang, T. J., & Perron, P. (1998). Additional tests for a unit root allowing for a break in the trend function at an unknown time. *International Economic Review*, 39, 1073-1100.
- Zivot, E., & Andrews, D. W. K. (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.

Rodney Fort is a professor of economic sciences at Washington State University.

Young Hoon Lee is a professor of economics at Hansung University.