

CUSUM Stability Testing of Estimated ARIMA Models¹

Theodore Bos and Thomas A. Fetherston

In the past 20 years, time series analysis has eclipsed regression analysis as the analysis of choice in the field of academic econometrics. The current time series paradigm was initiated by Box and Jenkins (1976) who synthesized a body of time series literature going back some 50 years. Their importance to the methodology is demonstrated by the fact the term *Box-Jenkins* is synonymous with *ARIMA*². Although ARIMA modeling is not unique to econometrics, the econometric need to explain the comovement of economic time series has led to the recent development of the uniquely econometric time series concept of *cointegration*³.

A necessary condition for the applicability of ARIMA modeling is that of *stationarity*. This condition enforces some regularity in the series' generating process, so that one might expect that the future looks somewhat like the past. Let the variable X_t denote a time series which is measured at fixed intervals over time. An ARIMA model can be written as

$$X_t - \phi_1 X_{t-1} - \dots - \phi_p X_{t-p} = C + a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q} \quad (1)$$

where a_t is white noise with a mean of zero and a constant variance. The series X_t is said to be *stationary* if its mean, $E(X_t)$, and its autocovariance, $Cov(X_t, X_s)$ for any integers t and s are constant over time. It is generally understood that this requirement prevents the direct use of trending data which is quite prevalent in economics and finance. This drawback is usually circumvented by the application of first differences, and very rarely second differences. Also the logarithms of the data are often taken prior to differencing. Over the years, the term *nonstationary* has come to mean that a series has some trend characteristics that require differencing. The ARIMA model can clearly be nonstationary by changes in specification. Several possibilities have been researched. One is where the constant C is postulated to change at a known point in time⁴. Another is to allow the parameters of the ARIMA model to change smoothly over time⁵. The common factor in such nonstationary models is that the form of the evolutionary process of the parameters must be known. We argue that we cannot *a priori* know *all* changes to the generating processes, and thus since stationarity is required, it must be tested for. We therefore examine a test for unspecified changes in an ARIMA process.

Brown, Durbin, and Evans (BDE, 1975) introduced the use of CUSUM and CUSUM of squares tests for the stability (or stationarity) of regression equations. Although BDE's CUSUM test based on the quality control CUSUM, it is still a distant relative to it. BDE's tests are based on what they called recursive residuals, the series of standardized one-step-ahead forecasts calculated from an ever-growing data set. Thus the regression equation is estimated on a small subset of the data, say the first several observations, and the standardized first-step-ahead forecast is calculated; the next observation is added to the data set and the next recursive residual is calculated; etc. BDE proved mathematically that these recursive residuals are independent, a characteristic that ordinary regression residuals do not have. As the procedure developed by BDE was suggested as a yardstick, and not a test, for instability, Bos and Fetherston (BF, 1992) proposed a procedure that would remedy this matter.

For example, they created simple regression data in which the intercept has one value up to period 27 and another value afterwards. The CUSUM plot and critical limits from the BDE and BF procedures are shown in Figure 1. The BDE procedure uses the straight lines as boundaries. If CUSUM *crosses* or goes beyond such linear boundaries at least once, then it is said that the regression equation is unstable. BF (1992) showed that the use of the BDE straight line boundaries would lead to an inequitable distribution of type I errors. To circumvent this drawback, they proposed curved boundaries for both the CUSUM and CUSUM of squares tests. As the BF boundaries are closer in than the BDE ones, more random crossings are expected. Thus BF proposed a procedure in which the crossings beyond the curved boundaries are counted and compared to a *critical number of crossings*. The critical number of crossings was determined by simulation. BF (1995) eased the computational burden to the researcher by providing formulas by which the critical number of

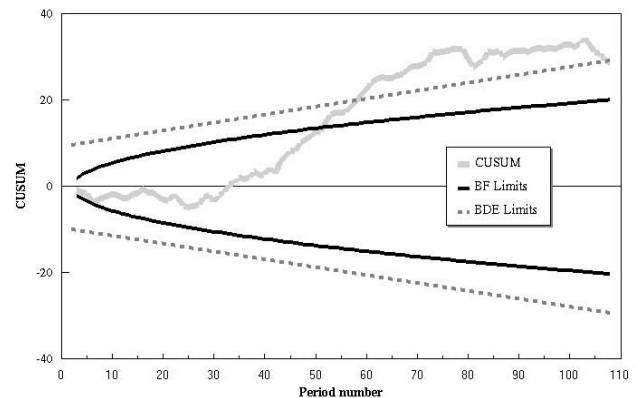


Figure 1. CUSUM Plot with BDE and BF Limits

crossings can be determined based on the number of recursive residuals used. Those formulas were derived by first seeing that the critical number of crossings is linear function of the number of recursive residuals; and then simulating histograms of crossings for situations with 10^2 , 10^3 , and 10^4 recursive residuals. More than 10^6 replications were done to accomplish this task.

On the basis of several propositions from earlier works, Smith (1985) asserted that the recursive residuals from estimated ARIMA models are also independent. This is not a far-fetched proposition as it is known that the estimate of an ARIMA(1,0,0) or AR(1) model is close to that of the corresponding regression model. We decided to check how well Smith's assertion would work in practice in view of the fact that ARIMA parameter estimates are not closed-form solutions as are regression estimates. We aim to see how tiny errors that will necessarily arise from the optimization procedures used in ARIMA estimation will affect the critical number of crossings in practice.

Because of the large number of possible models, including cumbersome seasonal models, we restrict this research to simple situations. The models we look at are:

(I) An AR(1) model estimated on X_t that follows an AR(1) process with $\theta_1 = -0.8$.

(II) An MA(1) model estimated on X_t that follows an MA(1) process with $\theta_1 = 0.5$.

(III) An AR(2) model estimated on X_t that follows an AR(1) process with $\theta_1 = 0.5$.

Models I and II are examples of simple properly specified ARIMA models, while model III is an example of an over-specified model. Over-specification of an ARIMA model is no problem, as the estimation can render the extra parameter to be essentially zero. The model I data was estimated by an approximate closed-form solution⁶. Models II and III were estimated using a quasi-Newton search procedure with function calls to likelihood function calculations carried out with Harvey's (1981) Kalman filter procedure.

In terms of model I, the AR(1) model on AR(1) data, our computer experiment was set up as follows:

(i) 100 AR(1) data points were generated.

(ii) Starting at observation 50, an AR(1) model was estimated.

(iii) The next observation was forecast, and the forecast error was recorded.

(iv) Steps (ii) and (iii) were repeated sequentially until the 100th observation was predicted.

(v) The AR(1) model was estimated on all 100 observations and the estimated error variance was recorded.

(vi) The 50 recorded forecast errors were divided by the square root of the final error variance to yield a stream of recursive residuals.

(vii) The recursive residuals, w_r were used to calculate the CUSUM, W_r , and CUSUM of squares, S_r , statistics, where

$$W_r = \sum_{t=51}^r w_t \quad (2)$$

$$\text{and } S_r = \frac{\sum_{t=51}^r w_t^2}{\sum_{t=51}^{100} w_t^2} \quad (3)$$

(viii) The number of crossings of the Bos-Fetherston boundaries were recorded.

(ix) Steps (i) through (viii) were repeated 1,000,000 times for model I, 100,000 times for model II, and 50,000 times for model III (different numbers of replications were done because of the computational expense involved with the search procedures and additional parameters). Thus the data for the histograms of the CUSUM and CUSUM of squares crossings were collected.

The resulting histograms are shown for the three cases in Figures 2a and 2b for CUSUM and CUSUM of squares, respectively. It is clear that the critical number of crossings is imperceptibly different from the crossings observed from simulated white noise from 10 million replications also shown on this figure. Table 1 presents the estimated critical number of crossings of the three models, the white noise described above, and the results from the formulas provided by BF (1995). It is clear from Figure 2 and Table 1 that there are differences among the situations presented and that these differences are due to numerical artifacts and not statistical ones. The scatter of

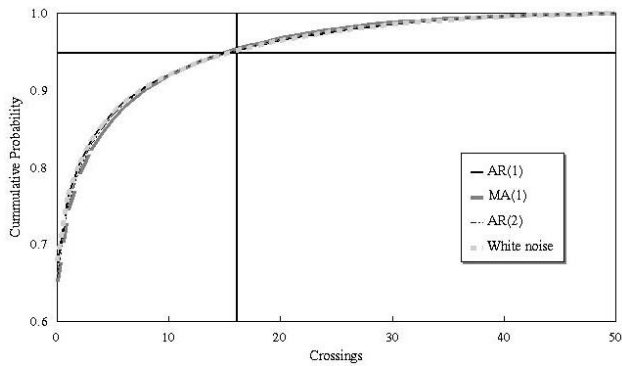


Figure 2a. CUSUM Cumulative Histograms

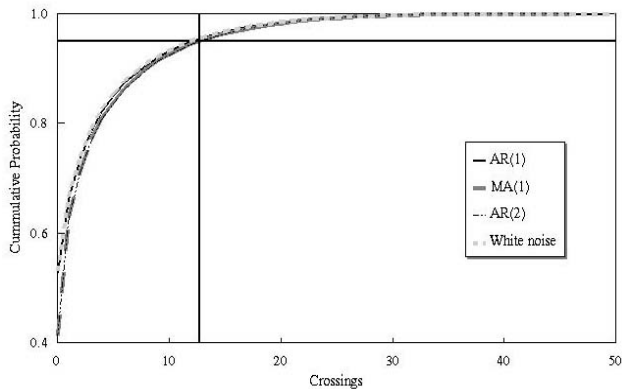


Figure 2b. CUSUM of Squares Cumulative Histograms

95th percentiles shows no obvious pattern due to errors from the search procedures or the approximated closed-form solutions used in model *I*. In a practical sense, since the BF (1995) values are larger than all others shown here, except for model *II* with CUSUM of squares, critical crossings generated from BF (1995) are generally safe.

TABLE 1. Estimated CUSUM and CUSUM of Squares Critical Crossings for Selected Models.

	Number of Replications	CUSUM	CUSUM of Squares
BF(1995)	--	15.996	12.652
White Noise	10 million	15.663	11.953
Case (a)	1 million	15.399	12.202
Case (b)	100,000	15.135	12.743
Case (c)	50,000	15.632	12.153

Box and Jenkins (1976, p. 293) considered the possibility of model inadequacy, which ... occurs when the form of the model remains the same but when the parameters change over a prolonged period of time. They discussed this in terms of Series B: IBM common stock closing prices: daily, 17 May 1961 through 2 November 1962. Earlier in the text, they found an ARIMA(0,1,1)

model to be adequate. In their section 8.2.3, they split the data in two halves and found an estimated θ of -0.29 for first half and -0.03 for the second half, while they found a value of -0.09 for the complete data set. For our analysis, we test the stability of the ARIMA(0,1,1) model. We



Figure 3a. CUSUM of an ARIMA(0,1,1) Model on Box-Jenkins Series B

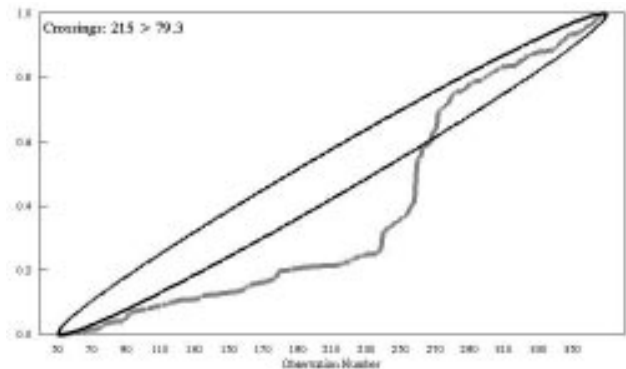


Figure 3b. CUSUM of an ARIMA(0,1,1) Model on Box-Jenkins Series B

start calculating recursive residuals after the 51st observation, thus not estimating ARIMA models with less than 50 observations. This yields us 319 recursive residuals. The number of CUSUM crossings is zero which is clearly less than the critical number of 101 crossings; the number of CUSUM of squares crossings is 215 which is much more than the 79.3 critical number of CUSUM of squares crossings with 319 recursive residuals. (These results are charted in Figures 3a and 3b.) Thus we find significant instability of the ARIMA(0,1,1) model with our CUSUM of squares test. This supports the quick analysis done by Box and Jenkins on both halves of the data set.

Acknowledgement

The authors would like to thank Eric Bradford and the staff of the AREN Helpdesk for their assistance with the computational part of this research.

References

Bos, Theodore and Thomas A. Fetherston, "Market Model Nonstationarity in the Korean Stock Market," *Pacific-Basin Capital Markets Research*, Volume III, ed. by G. Rhee and R. Chang. Amsterdam: North-Holland (1992).

Bos, T., and T. A. Fetherston. Nonstationarity of the Market Model, Outliers, and Choice of the Market Rate of Return. *Advances in Pacific Basin Financial Markets*, Volume 1, ed. by T. Bos and T. A. Fetherston, JAI Press Inc., Greenwich, Connecticut (1995).

Box, G. E. P., and G. M. Jenkins. (1976) *Time Series Analysis, Forecasting and Control* Holden-Day, San Francisco, CA, 1976.

Brown, R. L., J. Durbin, and J. M. Evans. (1975) Techniques for Testing the Constancy of Regression Relationships over Time. *JRSS(B)* 37:2, 149-163.

Engle, R. F., and C. W. J. Granger. (1987) Co-integration and Error Correction: Representation, Estimation, and Testing. *Econometrica*, 55(2):251-276.

Harvey, A. C., (1981) *Time Series Models* John Wiley & Sons, New York.

Smith, J. Q., (1985) Diagnostic Checks of Non-Standard Time Series Models. *Journal of Forecasting*, 4(3):283-291.

Endnotes

¹ This paper is based largely on the working paper by T. Bos and T. A. Fetherston. "CUSUM Testing for the Stability of ARIMA Estimated Models and for the Stationarity of Time Series Data."

² ARIMA stands for integrated autoregressive - moving average.

³ Engle, R. F., and C. W. J. Granger (1987)

⁴ See G. E. P. Box and G. C. Tiao. 1975. "Intervention Analysis with Applications to Economic and Environmental Problems" *JASA*, 70:70-79.

⁵ See C. Grillenzoni "ARIMA Processes With ARIMA Parameters" *Journal of Business & Economic Statistics*, April 1993, 11(2):235-250.

⁶ The algebra for this approximation can be obtained from the authors.

AREN Exhibit Wins an ADDY

AREN's Supercomputing '94 exhibit won a Citation of Excellence award in the trade show exhibit category in the Huntsville Advertising Federation's prestigious ADDY Awards Competition.

With guidance from AREN staff, the NRC Publications group developed an exhibit designed to achieve AREN marketing goals. Large formatted photographs which conveyed in-



formation on the facilities and capabilities of AREN were displayed on green two-toned walls accented by a gold band flowing through the entire exhibit. An interactive kiosk operated by touch screen allowed the viewer to stroll through ASA's interactive information center and surf the worldwide web with stops at ASA's or NRC's home pages.

The ADDY Awards Competition is the largest creative excellence competition designed to recognize outstanding advertising produced in the Huntsville area and nationwide. It is conducted annually by the American Advertising Federation. NRC Publications Division has been associated with the Huntsville Advertising Federation since 1990.

Credits:

Creative Director	Kevin Naumann, Cassandra Miller
Art Director	Kevin Naumann, Krel Shults
Copywriter	Sue DeLary
Photographer	Charlie Seifried
Illustrator	Bill Stem, Alan Jones, Krel Shults
Photo Productions	Quality Photo
Exhibit Productions	Industrial Graphics
Video Productions	Krel Shults

ERRATUM

In the February 1995 issue of *Interface*, Spring Hill College was inadvertently omitted from the statewide network diagram that appeared on the front page. Spring Hill College came on-line in January of 1995 with a connection to the University of South Alabama, which is the local AREN T1 access point. Glenn Bell, Director of Academic Computing, is the point of contact at Spring Hill College for questions about network access on that campus.