Econometrica, Vol. 60, No. 2 (March, 1992), 271–285

THE CUSUM TEST WITH OLS RESIDUALS[1]

BY WERNER PLOBERGER AND WALTER KRÄMER

We show that the CUSUM test of the stability over time of the coefficients of a linear regression model, which is usually based on recursive residuals, can also be applied to ordinary least squares residuals. We derive the limiting null distribution of the resulting test and compare its local power to that of the standard procedure. It turns out that neither version is uniformly superior to the other.

KEYWORDS: CUSUM test, least squares residuals.

1. INTRODUCTION AND SUMMARY

THIS PAPER FOLLOWS up the seminal work of Brown, Durbin, and Evans (1975; henceforth BDE) on the CUSUM and CUSUM of squares tests for the constancy over time of the coefficients of a linear regression model. Both tests are based on recursive residuals, which are independent $N(0, \sigma^2)$ under H_0, and therefore ideal ingredients for all types of tests. While the CUSUM of squares test has been generalized to ordinary least squares (OLS) residuals (McCabe and Harrison (1980)), a generalization of the CUSUM test to OLS residuals, which are dependent and heteroskedastic even under H_0, has not yet been made except for a very special case (MacNeill (1978)). In fact, both BDE (1975, p. 151) and McCabe and Harrison (1980, p. 142) argue that a least squares variant of the CUSUM test poses "intractable problems," due to an alleged difficulty in assessing the significance of the departure of the cumulated OLS residuals from their mean value zero.

Below we show that it is no more difficult to derive the limiting distribution for a CUSUM test based on OLS residuals than for a CUSUM test based on recursive residuals. While the CUSUMs of the recursive residuals, properly standardized, tend in distribution to a standard Wiener process (Sen (1982); Krämer, Ploberger, and Alt (1988), henceforth KPA), we show that the OLS-based CUSUMs tend in distribution to a Brownian bridge (or "tied-down-Brownian-motion;" see Billingsley (1968, p. 64)). This generalizes MacNeill's (1978) polynomial regression result to arbitrary regressor sequences. We also demonstrate that the resulting CUSUM test has higher (local) power for certain types of structural change than the one based on recursive residuals, though neither variant is uniformly superior to the other.

[1] This research was supported by Deutsche Forschungsgemeinschaft. We are grateful to Irene Schlüter for programming our Monte Carlo experiments, to Liqun Wang for checking the proofs of our theorems, and to Grayham Mizon, an editor, and two anonymous referees for many helpful insights and comments.

2. THE MODEL AND THE TEST

We consider the standard linear regression model

$$(1) \qquad y_t = x_t'\beta + u_t \qquad (t = 1, \ldots, T),$$

where at time t, y is the observation on the dependent variable, $x_t = [1, x_{t2}, \ldots, x_{tK}]' = [1, \tilde{x}_t']'$ is a $K \times 1$ vector of observations on the independent variables, with first component equal to unity, u are iid $(0, \sigma^2)$ disturbances (not necessarily normal), and β is the $K \times 1$ vector of regression coefficients. Below we are concerned with testing against the alternative that this unknown coefficient vector varies over time. Following BDE (1975), we do not specify a particular pattern of possible coefficient variation. We rather view the CUSUM procedures as pure significance tests, with power against various alternatives, some of which will be discussed below.

The null distribution of the local power of the tests are derived under the following assumptions:

Assumption A.1: *The regressors x_t and the disturbances u_t are defined on a common probability space, such that*

$$(2) \qquad \limsup_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} \|x_t\|^{2+\delta} < \infty \quad a.s.$$

for some $\delta > 0$ ($\|\cdot\|$ the Euclidean norm).

Assumption A.2:

$$(3) \qquad \lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} x_t x_t' = R \quad a.s.$$

for some nonsingular, nonstochastic $(K \times K)$ matrix R, where we assume without loss of generality that the model has been reparameterized such that

$$(4) \qquad R = \begin{bmatrix} 1 & 0 \\ 0 & R^* \end{bmatrix}.$$

In particular, this implies that

$$(5) \qquad c \equiv \lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} x_t = [1, 0, \ldots, 0]',$$

a fact we will repeatedly use below.

Assumption A.3: *The disturbances u_t are stationary and ergodic, with*

$$(6) \qquad E(u_t | \mathfrak{U}_t) = 0, \quad E(u_t^2 | \mathfrak{U}_t) = \sigma^2,$$

where $\mathfrak{U}_t$ is the σ-field generated by $\{y_{t-s}, x_{t-s+1}, u_{t-s} | s \geqslant 1\}$.

Assumptions A.1–A.3 conform to KPA (1988). They allow in particular for dynamic models, in which case they imply stability.

The standard CUSUM test is based on cumulated sums of the recursive residuals $\tilde{u}_t$ $(t = K+1, \dots, T)$. It rejects for large values of

$$\sup_{0 \leqslant z \leqslant 1} \left| \frac{W^{(T)}(z)}{1+2z} \right|, \tag{7}$$

where

$$W^{(T)}(z) = \frac{1}{\hat{\sigma}\sqrt{T}} \sum_{t=K+1}^{k+z(T-K)} \tilde{u}_t \quad (0 \leqslant z \leqslant 1), \tag{8}$$

and where $\hat{\sigma} = ((1/T)\Sigma_{t=k+1}^{T} \tilde{u}_t^2)^{1/2}$. This is equivalent to the initial rule by BDE (1975) to reject H_0 whenever $W_r^{(T)}(r) \equiv (1/\hat{\sigma})\Sigma_{t=k+1}^{r} \tilde{u}_t$ $(r = K+1, \dots, T)$ crosses either of the lines

$$a\sqrt{T-K} + 2a\frac{r-k}{\sqrt{T-K}} \quad \text{or} \quad -a\sqrt{T-K} - 2a\frac{r-k}{\sqrt{T-K}}, \tag{9}$$

where the parameter a depends on the significance level of the test. Arguing in terms of (8) however allows us to view the standardized sample paths $W^{(T)}(z)$ as random elements into the space $D[0,1]$ of all real valued functions on the $[0,1]$ interval that are right continuous and have left limits, i.e. to talk about convergence in a meaningful way. In particular it can then be shown that

$$W^{(T)}(z) \xrightarrow{\mathrm{d}} W(z) \quad \text{as} \quad T \mapsto \infty,$$

where "$\xrightarrow{\mathrm{d}}$" denotes convergence in distribution, and where $W(z)$ is the standard Wiener Process (Sen (1982), KPA (1988)). The appropriate critical values for the test statistic (7), given some significance level α, are then deduced from well known boundary crossing probabilities of the Wiener process.

This paper is concerned with the analog of (8) when the recursive residuals $\tilde{u}_t$ are replaced by OLS residuals

$$\hat{u}_t^{(T)} = y_t - x_t'\hat{\beta}^{(T)},$$

where $\hat{\beta}^{(T)} = (\Sigma_{t=1}^{T} x_t x_t')^{-1} \Sigma_{t=1}^{T} x_t y_t$ in the OLS coefficient estimate and where the superscript T emphasizes the dependence of these quantities on sample size. Unlike recursive residuals, OLS residuals usually change whenever another observation is added to the sample. In addition they are correlated and heteroskedastic even under H_0, so the null distribution of any test using them is much harder to derive. The main rationale for preferring them is that they better approximate the true disturbances under H_0, so any model deviations affecting the disturbances might be easier to detect.

An additional complication in the present context is that OLS residuals sum to zero when there is an intercept in the regression, as we have assumed. One therefore cannot expect their cumulated sum to drift off after a structural change, as often happens with recursive residuals, and which provides the rationale for the standard CUSUM test. No matter how large a structural shift has occurred, the cumulated OLS residuals will eventually return to the origin. The critical lines (9), with positive and negative slope, respectively, are therefore not appropriate with OLS residuals.

A natural alternative would be critical lines parallel to the horizontal axis, i.e., to reject the null hypothesis of parameter constancy whenever the maximum cumulated sum of OLS residuals becomes too large in absolute value. Other functions of the CUSUM sample paths, such as the L_2-norm, could be considered as well, and might even be more powerful, but we have not been able to derive the null distribution of the resulting test.

3. LIMITING DISTRIBUTION UNDER H_0

Let

$$B^{(T)}(z) = \frac{1}{\hat{\sigma}\sqrt{T}} \sum_{t=1}^{Tz} \hat{u}_t^{(T)}$$

denote the cumulated sums of the OLS residuals. Similar to the standard version of the CUSUM test, we have redefined the sample paths to be random elements from the underlying probability space into $D[0,1]$. The test statistic implied by our rejection rule is

$$\sup_{0 \leqslant z \leqslant 1} |B^{(T)}(z)|. \tag{10}$$

THEOREM 1: *Under H_0, and given Assumptions* A.1–A.3,

$$B^{(T)}(z) \xrightarrow{d} B(z) \quad as \quad T \to \infty, \tag{11}$$

where $B(z)$ is the standard Brownian bridge.

PROOF: Rewrite $\hat{u}_t^{(T)}$ as

$$\hat{u}_t^{(T)} = u_t - x_t'(\hat{\beta}^{(T)} - \beta)$$

and consider

$$\hat{\sigma} B^{(T)}(z) = \frac{1}{\sqrt{T}} \sum_{t=1}^{Tz} \hat{u}_t^{(T)} = \frac{1}{\sqrt{T}} \sum_{t=1}^{Tz} u_t - \frac{1}{\sqrt{T}} \sum_{t=1}^{Tz} x_t'(\hat{\beta}^{(T)} - \beta). \tag{12}$$

The relationship (11) then follows from

$$\operatorname*{plim}_{T \mapsto \infty} \sup_{0 \leqslant z \leqslant 1} \left| \frac{1}{\sqrt{T}} \left(\sum_{t=1}^{Tz} x_t'(\hat{\beta}^{(T)} - \beta) - z \sum_{t=1}^{T} u_t \right) \right| = 0 \tag{13}$$

(uniformly in z) and the well known fact that under Assumption A.3

$$\frac{1}{\sqrt{T}} \left(\sum_{t=1}^{Tz} u_t - z \sum_{t=1}^{T} u_t \right) \xrightarrow{d} \sigma B(z) \tag{14}$$

(see, e.g., McLeish (1975)). For proof of (13), consider

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{Tz} x_t'\left(\hat{\beta}^{(T)}-\beta\right)=\left[\frac{1}{T}\sum_{t=1}^{Tz} x_t'\right]\cdot\left[\sqrt{T}\left(\hat{\beta}^{(T)}-\beta\right)\right], \tag{15}$$

where the first term on the right-hand side tends in view of (5) to

$$zc' = [z, 0, \ldots, 0] \tag{16}$$

as $T \to \infty$, and where the second term can, because of (4), be expressed as

$$\sqrt{T}\left(\hat{\beta}^{(T)}-\beta\right)=\frac{1}{\sqrt{T}}\begin{bmatrix}1 & 0\\ 0 & R^*\end{bmatrix}^{-1}\begin{bmatrix}\sum_{t=1}^{T} u_t\\ \sum_{t=1}^{T} \tilde{x}_t u_t\end{bmatrix}+o_p(1). \tag{17}$$

The relationships (16) and (17) imply that

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{Tz} x_t'\left(\hat{\beta}^{(T)}-\beta\right)=\frac{z}{\sqrt{T}}\sum_{t=1}^{T} u_t + o_p(1) \quad \text{(uniformly in } z\text{)},$$

which proves (13) and, in conjunction with (14) and $\operatorname{plim}_{T\mapsto\infty}\hat{\sigma}=\sigma$, the theorem. *Q.E.D.*

Theorem 1 immediately yields the limiting distribution of the test statistic (10) itself. Since $\sup_{0\leqslant z\leqslant 1}|B^{(T)}(z)|$ is a continuous function of $B^{(T)}$ (where $D[0,1]$ is endowed with the Skorohod topology), we have

$$\sup_{0\leqslant z\leqslant 1}\left|B^{(T)}(z)\right| \xrightarrow{d} \sup_{0\leqslant z\leqslant 1}\left|B(z)\right|. \tag{18}$$

Since, in addition, the event

$$\left\{\sup_{0\leqslant z\leqslant 1}\left|B^{(T)}(z)\right|=a\right\}$$

has B-measure zero for all real a, this implies that

$$P\left(\sup_{0\leqslant z\leqslant 1}\left|B^{(T)}(z)\right|>a\right)\to P\left(\sup_{0\leqslant z\leqslant 1}\left|B(z)\right|>a\right) \tag{19}$$

as $T \to \infty$, where the latter probability is well known (see, e.g., Billingsley (1968, p. 85)) and equal to

$$2\sum_{j=1}^{\infty}(-1)^{j+1}\exp\left(-2j^2a^2\right). \tag{20}$$

The limiting distribution function of the test statistic (10) is therefore one minus the expression (20). This is identical to the limiting null distribution (for $K=1$) of the Ploberger-Krämer-Kontrus fluctuation test (1989) and also of the CUSUM test with mean adjusted data (Ploberger and Krämer (1987)). Some useful critical values are 1.22 ($\alpha = 10\%$), 1.36 ($\alpha = 5\%$), and 1.63 ($\alpha = 1\%$).

The variance of a Brownian bridge varies across the [0, 1] interval and is given by

$$\text{(21)} \qquad \operatorname{var}(B(z)) = z(1-z).$$

This implies that the rejection probability under H_0 is likewise varying across [0, 1], in the sense that the probability that the sample path $B^{(T)}(z)$ crosses a critical line parallel to the z-axis is at a maximum for $z = 1/2$. We therefore conjecture that the power of the test can be improved by critical curves which spread the rejection probability under H_0 more evenly across the [0, 1]-interval. As in BDE (1975), our particular choice was dictated by mathematical convenience rather than an optimization argument, since crossing probabilities for curves other than straight lines are notoriously hard to find.

Similar to the standard version, the above OLS-based CUSUM test attains its nominal size only asymptotically. In addition, and unlike the standard version with normal disturbances, its null distribution also varies across design matrices. However, we provide some Monte Carlo evidence below that the asymptotic approximation works well for moderate sample sizes, and that the test is almost always conservative.

4. LOCAL POWER

Next we let the regression coefficients in (1) vary according to

$$\text{(22)} \qquad \beta_{t,T} = \beta + \frac{1}{\sqrt{T}} g(t/T),$$

where $g(z)$ is an arbitrary K-dimensional function defined on the [0, 1] interval. Following KPA (1988), we only require that g can be expressed as a uniform limit of functions that are constant on intervals (which implies that g and all its components are bounded on the [0, 1] interval). When $g(z) = 0$ $(z < z^*)$ and $g(z) = \Delta\beta (z \geqslant z^*)$, (22) includes a one-time shift of the regression coefficients at time $T^* = z^*T$ as a special case, the intensity of which decreases with sample size and which is always located at the same quantile of the sample observations.

Given alternatives (22), Ploberger and Krämer (1990) and KPA (1988) show that

$$\text{(23)} \qquad W^{(T)}(z) \xrightarrow{d} W(z) + \frac{1}{\sigma}\left[c'\int_0^z g(v)\,dv - c'\int_0^z \left(\frac{1}{v}\int_0^v g(w)\,dw\right) dv\right],$$

i.e. that the standard CUSUM sample paths tend in distribution to a Wiener process plus some nonstochastic function, which depends on the structural change and the mean regressor c. In particular this nonstochastic function is identically equal to zero if all structural shifts are orthogonal to c, in which case the standard CUSUM test has local power equal to its size.

Next we derive the analogue of (23) for the cumulated OLS residuals $B^{(T)}(z)$. For this purpose, let

$$\tilde{y}_{t,T} = x_t'\beta_{t,T} + u_t \qquad (t = 1,\ldots,T) \tag{24}$$

be the triangle sequence of dependent variables under the local alternatives (22). Putting the model like this implicitly excludes lagged dependent variables on the right-hand side, which under the alternative would imply a triangle sequence rather than a simple sequence of regressors x_t as well. This would greatly complicate notation, without affecting the results, so we confine our formal analysis to the nondynamic case. However, in the Appendix we prove that the exclusion of lagged dependent variables does not matter here.

Let

$$\tilde{\beta}^{(T)} = \left(\sum_{t=1}^{T} x_t x_t'\right)^{-1} \sum_{t=1}^{T} x_t' \tilde{y}_{t,T} \tag{25}$$

be the OLS coefficient estimator for a given sample size T.

THEOREM 2: *Under Assumptions* A.1–A.3 *and given local alternatives* (22),

$$B^{(T)}(z) \xrightarrow{d} B(z) + \frac{1}{\sigma}\left[\int_0^z c'g(u)\,du - c'z\int_0^1 g(u)\,du\right]. \tag{26}$$

PROOF: We have

$$\tilde{\beta}^{(T)} = \hat{\beta}^{(T)} + \frac{1}{\sqrt{T}} d_T, \tag{27}$$

where

$$d_T = \left(\sum_{t=1}^{T} x_t x_t'\right)^{-1} \left(\sum_{t=1}^{T} x_t x_t' g\left(\frac{t}{T}\right)\right), \tag{28}$$

so

$$\begin{aligned} \tilde{y}_{t,T} - x_t'\tilde{\beta}^{(T)} &= u_t + x_t'\beta_{t,T} - x_t'\tilde{\beta}^{(T)} \\ &= u_t + x_t'\beta + x_t' g\left(\frac{t}{T}\right)\frac{1}{\sqrt{T}} - x_t'\hat{\beta}^{(T)} - x_t' d_T \frac{1}{\sqrt{T}} \\ &= y_t - x_t'\hat{\beta}^{(T)} + x_t' g\left(\frac{t}{T}\right)\frac{1}{\sqrt{T}} - x_t' d_T \frac{1}{\sqrt{T}} \\ &= \hat{u}_t^{(T)} + x_t' g\left(\frac{t}{T}\right)\frac{1}{\sqrt{T}} - x_t' d_T \frac{1}{\sqrt{T}}. \end{aligned} \tag{29}$$

This implies that

$$\text{(30)} \qquad \frac{1}{\sqrt{T}} \sum_{t=1}^{Tz} \left(\tilde{y}_{t,T} - x_t' \tilde{\beta}^{(T)} \right)$$
$$= \frac{1}{\sqrt{T}} \sum_{t=1}^{Tz} \hat{u}_t^{(T)} + \frac{1}{T} \sum_{t=1}^{Tz} x_t' g(t/T) - \frac{1}{T} \sum_{t=1}^{Tz} x_t' d_T,$$

where the first term tends in distribution to $\sigma B(z)$ by Section 3. By Lemma 4 in KPA (1988), the second term tends, uniformly in z, to $\int_0^z c' g(u)\, du$. As to the last term, note that $(1/T)\Sigma_{t=1}^{Tz} x_t$ tends to $c'z$ (uniformly in z), and that

$$\text{(31)} \qquad d_T = \left(\frac{1}{T} \sum_{t=1}^{T} x_t x_t' \right)^{-1} \frac{1}{T} \sum_{t=1}^{T} x_t x_t' g\left(\frac{t}{T} \right) \to R^{-1} R \int_0^1 g(u)\, du = \int_0^1 g(u)\, du.$$

Together, this implies that, uniformly in z,

$$\text{(32)} \qquad \frac{1}{T} \sum_{t=1}^{Tz} x_t' d_T \to c'z \int_0^1 g(u)\, du.$$

Since in addition $\hat{\sigma}$ is easily shown to remain consistent under local alternatives (22), this completes the proof of the theorem. *Q.E.D.*

From the expression (26) it is obvious that the OLS based CUSUM test has only trivial local power against structural changes that are orthogonal to the mean regressor c. Similarly to the standard version, the CUSUM sample paths have then the same limiting distribution as under H_0.

If $g(z)$ is not orthogonal to c for all z in $[0,1]$, the limiting rejection probabilities exceed the size for both versions of the test and could, in principle, be calculated as boundary crossing probabilities of the stochastic processes in (23) and (26). However, the practical evaluation of these probabilities appears intractable even for very simple g functions. We hope to report more definite results on this technical problem in a later paper.

Some general conclusions can still be drawn. First and quite expectedly, the structural change and the disturbance standard deviation enter the respective local power expressions only via $g(z)/\sigma$. This means that large shifts in conjunction with large disturbances have asymptotically the same chance of being detected as small shifts in a model with small disturbances. Second, if there is a single shift at time $T^* = z^*T$ (i.e. $g(z) = 0$ for $z < z^*$ and $g(z) = \Delta\beta$ for $z \geqslant z^*$), then the right-hand side of (23) is a Wiener process up to $z = z^*$ and has a different distribution only for the remaining z. This means that a shift late in the sample is likely to go unnoticed by the standard CUSUM test. The right-hand side of (26), on the other hand, is affected for all z in $[0,1]$, no matter where the structural change occurs. Third, a single shift at time $T^* = z^*T$ induces the same limiting distribution of the OLS-based test statistic (10) as a shift of the same size at time $(1 - z^*)T$. Finally, and most importantly, structural changes affect the limiting rejection probabilities only via their inner product with the mean regressor c. This applies to both versions of the test, and renders their comparison much easier.

TABLE I

LOCAL POWER AGAINST SINGLE STRUCTURAL SHIFT ($\alpha = 5\%$)

q	Break Point z^* 0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
	(a) Standard CUSUM Test								
2	5.9	6.9	6.9	6.3	5.7	4.9	4.5	4.2	4.0
4	13.7	17.9	18.2	16.6	13.6	9.6	7.1	5.0	4.2
6	26.5	37.6	39.3	35.9	29.7	21.2	12.9	7.3	4.4
8	44.0	61.1	65.4	63.1	54.5	40.0	24.1	11.1	5.1
10	62.2	81.7	86.4	85.1	77.9	62.9	40.2	17.7	6.0
	(b) OLS-based CUSUM Test								
2	4.5	6.7	9.3	11.1	12.6	11.1	8.7	6.3	4.4
4	6.7	15.9	28.9	37.3	40.2	37.2	28.7	16.1	6.4
6	10.7	37.0	60.6	72.4	75.5	72.2	60.1	36.6	10.5
8	18.6	64.1	86.9	93.8	94.9	93.5	86.2	63.4	18.3
10	30.5	86.3	97.2	99.3	99.7	97.3	97.7	85.1	30.9

As an illustration, Table I gives a Monte Carlo approximation of their local power against a single structural shift. In this case, only two parameters enter the respective expressions for the limiting rejection probabilities. They are z^* (the timing of the shift) and q, which is determined by

$$\frac{1}{\sigma}c'g(z) = \begin{cases} 0, & z \leqslant z^*, \\ q, & z \geqslant z^*. \end{cases}$$

The Monte Carlo approximation to the local power was obtained by discrete readings (spaced $1/200$ apart) of the stochastic processes on the right-hand side of (23) and (26), respectively. With a single structural shift, the nonstochastic components of the processes are given by

$$(33) \qquad \frac{1}{\sigma}\left[c'\int_0^z g(v)\,dv - c'\int_0^z\left(\frac{1}{v}\int_0^v g(w)\,dw\right)\right]$$
$$= \begin{cases} 0, & z < z^*, \\ gz^*[\ln(z) - \ln(z^*)], & z \geqslant z^*, \end{cases}$$

and

$$(34) \qquad \frac{1}{\sigma}\left[\int_0^z c'g(u)\,du - c'z\int_0^1 g(u)\,du\right]$$
$$= \begin{cases} qz(z^* - 1), & z < z^*, \\ qz(z^* - 1) - q(z^* - z), & z \geqslant z^*, \end{cases}$$

respectively. $N = 5000$ runs were performed for any given q and z^*.

The table shows that the OLS-based version has almost uniformly better local power than the standard test. Its local power peaks for shifts at $z^* = 1/2$, and is symmetric about this point, as predicted by formula (26). (Differences in empirical rejection probabilities between shifts at z^* and $1 - z^*$ are due to sampling error.) The standard CUSUM test does best for shifts at around $z^* = 0.3$, confirming previous Monte Carlo work (e.g., KPA (1988), Ploberger et al. (1989), or Krämer et al. (1990)), and performs rather poorly as $z^* \to 1$. Only for very small z^* does it outperform the OLS-based CUSUM test.

5. RELATIVE POWER IN FINITE SAMPLES

Next we compare the finite sample power of the tests in another sampling experiment. For ease of comparison with similar experiments in KPA (1988) we confine ourselves to a single structural shift in a simple model where $K = 2$, $x_t = [1, (-1)^t]$, and $u_t \sim \text{nid}(0, 1)$. Additional Monte Carlo results are in Krämer et al. (1990).

We systematically vary (i) the timing, (ii) the intensity, and (iii) the angle of the shift, which is given by

$$(35) \qquad \Delta\beta = \frac{b}{\sqrt{T}}[\cos\psi, \sin\psi],$$

where ψ is the angle between $\Delta\beta$ and the mean regressor $c = [1, 0]'$. The intensity of the shift is $\|\Delta\beta\| = |b|/\sqrt{T}$. It occurs at time $T^* = z^*T$, with $z^* = 0.1$, 0.3, 0.5, 0.7, and 0.9. This covers shifts early, midway, and late in the sample period. $N = 1000$ runs were performed for any given parameter combination. Table II reports some representative results for $T = 120$ and $\alpha = 5\%$.

The table confirms the local power results from Section 4. Both tests perform worse as structural shift and mean regressor approach orthogonality. For $\psi = 90°$, power for both tests is even less than size. The standard CUSUM test outperforms the OLS-based version when the shift occurs rather early ($z^* = 0.1$ or $z^* = 0.3$), but gets beaten as z^* increases. Both tests improve as the intensity of the shift increases, except for high values of φ, and the OLS-based CUSUM test has similar power against shifts at z^* and $1 - z^*$.

We also simulated the null distribution of both tests for both regressor sequences, to correct for possible differences in size. With $N = 10.000$ runs and for $x_t = [1, (-1)^t]$, we obtained a true size of 3.77% for the standard CUSUM test and 3.78% for the OLS-based version. Both tests are therefore conservative, which for the standard CUSUM test has long been known, and have almost identical true size for the particular regressor sequence in our Monte Carlo experiment. Since differences between nominal and actual size are identical for both tests, there is no need to correct for this size distortion.

6. CONCLUSION

The CUSUM test can equally well be based on OLS residuals. It then reacts also to structural shifts which occur late in the sample, which are likely to go unnoticed by the standard version of the test. No version is uniformly superior to the other.

TABLE II

FINITE SAMPLE POWER OF THE CUSUM TESTS FOR $x_t = [1, (-1)^t]$

	b	Angle ψ 0°	18°	36°	54°	72°	90°
		(a) Standard CUSUM Test					
$z^* = 0.1$	4.8	.173	.153	.115	.075	.049	.029
	7.2	.372	.323	.228	.125	.058	.026
	9.6	.590	.544	.404	.200	.066	.023
	12.0	.777	.729	.570	.299	.087	.019
$z^* = 0.3$	4.8	.236	.210	.142	.084	.046	.025
	7.2	.546	.495	.336	.151	.057	.019
	9.6	.811	.752	.590	.271	.070	.013
	12.0	.949	.918	.777	.416	.081	.011
$z^* = 0.5$	4.8	.160	.137	.100	.071	.043	.027
	7.2	.412	.356	.231	.098	.044	.017
	9.6	.697	.620	.434	.169	.048	.013
	12.0	.884	.836	.645	.272	.048	.009
$z^* = 0.7$	4.8	.071	.069	.059	.046	.032	.028
	7.2	.144	.124	.089	.057	.029	.019
	9.6	.294	.245	.150	.064	.028	.015
	12.0	.490	.430	.247	.087	.026	.010
$z^* = 0.9$	4.8	.033	.032	.033	.033	.032	.032
	7.2	.039	.039	.035	.032	.030	.030
	9.6	.045	.043	.038	.029	.028	.026
	12.0	.049	.048	.040	.032	.022	.020
		(b) OLS-based CUSUM Test					
$z^* = 0.1$	4.8	.079	.075	.066	.048	.038	.038
	7.2	.130	.121	.093	.067	.039	.032
	9.6	.221	.199	.141	.084	.043	.028
	12.0	.391	.337	.211	.104	.043	.022
$z^* = 0.3$	4.8	.390	.346	.257	.143	.063	.033
	7.2	.779	.724	.538	.276	.083	.026
	9.6	.963	.939	.824	.464	.117	.020
	12.0	.999	.997	.954	.673	.153	.012
$z^* = 0.5$	4.8	.550	.505	.368	.203	.070	.032
	7.2	.906	.875	.719	.400	.114	.021
	9.6	.993	.985	.927	.633	.169	.016
	12.0	1.000	.999	.992	.830	.223	.005
$z^* = 0.7$	4.8	.405	.367	.256	.148	.058	.032
	7.2	.768	.714	.553	.272	.079	.021
	9.6	.965	.964	.811	.474	.120	.014
	12.0	.997	.994	.957	.668	.153	.008
$z^* = 0.9$	4.8	.074	.072	.058	.046	.038	.038
	7.2	.136	.123	.089	.059	.039	.032
	9.6	.240	.205	.144	.079	.038	.029
	12.0	.384	.338	.219	.103	.040	.025

Cowles Foundation for Research in Economics, Yale University, P.O. Box 2125 Yale Station, New Haven, CT 06520-2125, U.S.A., and Institut für Ökonometrie und Systemtheorie, TU Wien, A-1040 Wien, Austria

and

Universität Dortmund, FB Statistik, Lehrstuhl Wirtschafts- und Sozialstatistik, Vogelpothsweg 87, Postfach 50 05 00, D-4600, Dortmund 50, Germany

Manuscript received October, 1989; final revision received June, 1991.

APPENDIX

Proof of Theorem 2 in Dynamic Models

Let $x_t = [1, y_{t-1}, \dots, y_{t-q}, x_{t,q+2}, \dots, x_{t,K}]'$ be the regressors in a dynamic model, where we now explicitly distinguish in notation between lagged dependent and "truly" exogenous variables. With regression coefficients varying according to (22), we have, under the alternative, that the regressors

$$x_{t,T} = \left[1, y_{t-1,T}, \dots, y_{t-q,T}, x_{t,q+2}, \dots, x_{t,K}\right]'$$

are likewise a triangle sequence rather than a simple sequence, which is why things become much more complicated here. In order not to overburden the notation, we use similar symbols for the $K \times 1$ vector of right-hand-side variables at time t $(x_{t,T})$ and its elements, which meaning is however always clear from the context.

The key in generalizing Theorem 2 to dynamic models is the following result:

Lemma 1: *Under the assumptions of Theorem* 2, *we have*

(A1) $$\sum_{t=1}^{T} \|x_{t,T} - x_t\|^2 = O_p(1) \quad \text{and}$$

(A2) $$\sum_{t=1}^{T} (y_{t,T} - y_t)^2 = O_p(1).$$

Proof: Since (A1) immediately follows from (A2), it suffices to prove (A2). To this purpose, decompose $y_{t,T} - y_t$ into

(A3) $$y_{t,T} - y_t = \Delta_{t,T} + \delta_{t,T}, \quad \text{where}$$

$$\Delta_{t,T} = y_{t,T} - y^*_{t,T},$$

(A4) $$\delta_{t,T} = y^*_{t,T} - y_t,$$

(A5) $$y^*_{t,T} = \beta_1 + \frac{1}{\sqrt{T}} g_1(t/T) + \sum_{i=1}^{q} \beta_{i+1} y^*_{t-i,T} + \sum_{i=q+2}^{K} \left(\beta_i + \frac{1}{\sqrt{T}} g_i(t/T)\right) x_{t,i} + u_t,$$

and where, for $t \leqslant 0$,

(A6) $$y^*_{t,T} = y_{t,T} = y_t.$$

The $y^*_{t,T}$ are a modified triangle sequence of dependent variables where only the truly exogenous regressors are disturbed and g_i is the ith component of the K-dimensional function g. The proof of the Lemma now rests on certain bounds for both

$$\sum_{t=1}^{T} \delta^2_{t,T} \quad \text{and} \quad \sum_{t=1}^{T} \Delta^2_{t,T}.$$

To see this, decompose $\delta_{t,T}$ into

$$\delta_{t,T} = \sum_{i=1}^{q} \beta_{i+1}\delta_{t-i,T} + \xi_{t,T} \tag{A7}$$

where

$$\xi_{t,T} := \frac{1}{\sqrt{T}} g_1(t/T) + \frac{1}{\sqrt{T}} \sum_{i=q+2}^{K} g_i(t/T)x_{t,i} \tag{A8}$$

when $t > 0$ and $\xi_{t,T} := 0$ when $t \leqslant 0$ and where $\limsup \Sigma_{t=1}^T \xi^2_{t,T} < \infty$ in view of Assumption A.2 (equation 3) and the boundedness of g. In addition, because of the stability of the recursion (A7), there exists a sequence of constants γ_i, with $|\gamma_i| \leqslant M\rho^i$, $0 < \rho < 1$, such that

$$\delta_{t,T} = \sum_{i=0}^{t} \gamma_i \xi_{t-i,T}, \tag{A9}$$

which implies that

$$\begin{aligned} \sum_{t=1}^{T} \delta^2_{t,T} &= \sum_{t=1}^{T} \left[\sum_{i,j=0}^{t} \gamma_i \gamma_j \xi_{t-i,T}\xi_{t-j,T} \right] \\ &\leqslant \sum_{i,j=0}^{T} |\gamma_i||\gamma_j| \left| \sum_{t=1}^{T} \xi_{t-i,T}\xi_{t-j,T} \right| \\ &\leqslant \left[\sum_{i=0}^{T} \gamma_i^2 \right] \left[\sum_{t=1}^{T} \xi^2_{t,T} \right] \\ &\leqslant \text{some random variable } V_1. \end{aligned} \tag{A10}$$

In a similar fashion, we find a bound for $\Sigma_{t=1}^T \Delta^2_{t,T}$, by decomposing $\Delta_{t,T}$ into

$$\Delta_{t,T} = \sum_{i=1}^{q} \beta_{i+1} \Delta_{t-i,T} + \eta_{t,T} \tag{A11}$$

where

$$\eta_{t,T} := \frac{1}{\sqrt{T}} \sum_{i=1}^{q} g_i(t/T) y_{t-i,T} \tag{A12}$$

when $t > 0$ and $\eta_{t,T} = 0$ when $t < 0$. On the analogy to (A10), it is easily seen that

$$\sum_{t=1}^{T} \Delta^2_{t,T} \leqslant C\left(\sum_{t=1}^{T} \eta^2_{t,T} \right) \quad \text{for some constant } C. \tag{A13}$$

Also, by the boundedness of g,

$$\begin{aligned} \sum_{t=1}^{T} \eta^2_{t,T} &\leqslant C_2 \left[\frac{1}{T} \sum_{t=1}^{T} y^2_{t,T} + \frac{1}{T} \sum_{t=-q}^{0} y^2_{t,T} \right] \\ &\leqslant \left[\sqrt{C_2/T} \sum_{t=1}^{T} y^2_{t,T} + \sqrt{C_1} \right]^2 \end{aligned}$$

for some constant C_2 and some random variable $C_1 \geqq (1/T)\Sigma_{t=-q}^{0} y_{t,T}^2 = (1/T)\Sigma_{t=-q}^{0} y_t^2$. Then with $V_1' = C_1 C$ and $V_2 = C_2 C$, (A13) implies that

$$\text{(A14)} \qquad \sqrt{\sum_{t=1}^{T} \Delta_{t,T}^2} \leqslant \sqrt{V_2/T}\sqrt{\sum_{t=1}^{T} y_{t,T}^2} + \sqrt{V_1}\,.$$

From

$$y_{t,T} - y_t = \Delta_{t,T} + \delta_{t,T},$$

we therefore have

$$\text{(A15)} \qquad \sum_{t=1}^{T} (y_{t,T} - y_t)^2 \leqslant \sqrt{V_1} + \sqrt{V_1'} + \sqrt{V_2/T}\sqrt{\sum_{t=1}^{T} y_{t,T}^2}$$

and

$$\text{(A16)} \qquad \sqrt{\sum y_{t,T}^2} \leqslant \sqrt{\sum (y_{t,T} - y_t)^2} + \sqrt{\sum y_t^2}$$

$$\leqslant \sqrt{V_1} + \sqrt{V_1'} + \sqrt{V_2/T}\sqrt{\sum y_{t,T}^2} + \sqrt{T}\sqrt{V_3}\,,$$

where V_3 is some random variable that dominates $\Sigma y_t^2/T$. Solving for $(\Sigma y_{t,T}^2)^{1/2}$ yields for $T > V_2$:

$$\text{(A17)} \qquad \sqrt{\sum y_{t,T}^2} \leqslant \left(\sqrt{V_1} + \sqrt{V_1'} + \sqrt{V_3}\sqrt{T}\right) \Big/ \left(1 - \sqrt{V_2/T}\right),$$

and plugging this bound into (A15) proves the lemma. *Q.E.D.*

Now the remainder of the general version of Theorem 2 follows easily from (A1) and (A2). In the general case, expression (30) takes the form

$$\text{(A18)} \qquad \frac{1}{\sqrt{T}} \sum_{t=1}^{Tz} \left(\tilde{y}_{t,T} - x_{t,T}\tilde{\beta}^{(T)} \right)$$

$$= \frac{1}{\sqrt{T}} \sum_{t=1}^{Tz} u_t^{(T)} + \frac{1}{T} \sum_{t=1}^{Tz} x_{t,T}' g(t/T) - \frac{1}{T} \sum_{t=1}^{Tz} x_t' d_T,$$

where d_T is now different from (28) and given by

$$\text{(A19)} \qquad d_T = \left(\sum_{t=1}^{T} x_{t,T} x_{t,T}' \right)^{-1} \left(\sum_{t=1}^{T} x_{t,T} x_{t,T}' g(t/T) \right).$$

The first term in (A13) again tends in distribution to $\sigma B(z)$ by Section 3. The second term can be written as

$$\text{(A20)} \qquad \frac{1}{T} \sum_{t=1}^{Tz} x_t' g(t/T) + \frac{1}{T} \sum_{t=1}^{Tz} (x_{t,T} - x_t) g(t/T),$$

where the first term tends to $\int_0^z c' g(u)\, du$ and the second term vanishes (uniformly in z) in view of Lemma 1 and the boundedness of g. Along similar lines, by decomposing the term into one part that is identical to the standard case and one part that vanishes uniformly in z, we also show that the third term in (A18) attains the limit in (32), so Theorem 2 holds for dynamic models as well.

REFERENCES

Brown, R. L., J. Durbin, and J. M. Evans (1975): "Techniques for Testing the Constancy of Regression Relationships Over Time," *Journal of the Royal Statistical Society*, B, 37, 149–163.

Billingsley, P. (1968): *Convergence of Probability Measures*. New York: Wiley.

Krämer, W., W. Ploberger, and R. Alt (1988): "Testing for Structural Change in Dynamic Models," *Econometrica*, 56, 1355–1369.

Krämer, W., W. Ploberger, and I. Schlüter (1990): "Recursive vs. OLS-Residuals in the CUSUM Test," to appear in *Economic Structural Change: Analysis and Forecasting*, ed. by P. Hackl and A. Westlund. Heidelberg: Springer.

MacNeill, I. B. (1978): "Properties of Sequences of Partial Sums of Polynomial Regression Residuals with Applications to Tests for Change of Regression at Unknown Times," *The Annals of Statistics*, 6, 422–433.

McCabe, B. P. M., and M. J. Harrison (1980): "Testing the Constancy of Regression Relationships Over Time Using Least Squares Residuals," *Journal of the Royal Statistical Society*, C (*Applied Statistics*), 29, 142–148.

McLeish, P. L. (1975): "Invariance Principles for Dependent Variables," *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 32, 165–178.

Ploberger, W., and W. Krämer (1987): "Mean Adjustment and the CUSUM Test Structural Change," *Economics Letters*, 25, 255–258.

——— (1990): "The Local Power of the CUSUM and CUSUM of Squares Tests," *Econometric Theory*, 6, 335–347.

Ploberger, W., W. Krämer, and K. Kontrus (1989): "A New Test for Structural Stability in the Linear Regression Model," *Journal of Econometrics*, 40, 307–318.

Sen, P. K. (1982): "Invariance Principles for Recursive Residuals," *The Annals of Statistics*, 10, 307–312.