NEUERE UNTERSUCHUNGEN ZUM CUSUM-TEST

Anwendung auf dynamische Modelle
und Einführung einer
modifizierten Varianzschätzung

RAIMUND ALT*

Forschungsbericht/
Research Memorandum No. 225

Dezember 1985

* Scholar am Institut für Höhere Studien, Abteilung Mathematische Methoden und Computerverfahren.

ZUSAMMENFASSUNG

In dieser Arbeit werden die Ergebnisse von Monte Carlo Experimenten praesentiert, die zur Untersuchung einiger spezieller Eigenschaften des CUSUM-Tests durchgefuehrt wurden. Die in zahlreichen Tabellen enthaltenen empirischen Resultate beziehen sich auf das Signifikanzniveau bzw. die Teststaerke (power) der zugrundeliegenden Testvarianten. Der erste Teil, in dem die Anwendung des CUSUM-Tests auf dynamische Modelle untersucht wird, liefert einen Vergleich zwischen dem gewoehnlichen CUSUM-Test und einem Test, der auf dem Vorschlag von Dufour (1982) beruht. Anschließend wird, diesmal bei statischen Modellen, eine modifizierte Varianzschaetzung eingefuehrt, die auf eine Anregung von Harvey (1975) zurueckgeht. Die Ergebnisse zeigen (bei kleinen Stichproben) eine deutliche Ueberlegenheit des modifizierten Tests.

SUMMARY

The paper presents results of Monte Carlo experiments, which were designed to reveal some special properties of the CUSUM test. Numerous tables contain empirical results on the significance level and power of some modifications of the usual CUSUM test. The first part of the paper explores the application of the CUSUM test in dynamic models. Here, the usual CUSUM test is compared with a test suggested by Dufour (1982). Then, following an idea of Harvey (1975), a modified variance estimator is introduced for static models. The superiority of the modified test is evident, especially in small samples.

INHALTSVERZEICHNIS

I. Der CUSUM-Test und seine Anwendung auf dynamische Modelle.

1. Einleitung.

Der von Brown, Durbin und Evans (1975) entwickelte CUSUM-Test dient zur Ueberpruefung der Konstanz der Regressionskoeffizienten im linearen Modell. Dabei galt bislang (u.a.) als Voraussetzung, daß saemtliche Regressoren nichtstochastisch sind. Somit waeren z.B. dynamische Modelle von der Anwendung dieses Tests ausgeschlossen. Durch die Arbeit von Kraemer, Ploberger, Alt (1985) ist nunmehr der Nachweis erbracht, daß auch bei dynamischen Modellen (zumindest bei genuegend großem Stichprobenumfang) die Anwendung des CUSUM-Tests gerechtfertigt ist.

Im Mittelpunkt dieses Kapitels stehen nun eine Reihe von Monte Carlo Experimenten, die im Rahmen der letztgenannten Arbeit durchgefuehrt wurden. Bevor die Ergebnisse dieser Experimente praesentiert werden, soll aber noch kurz auf die theoretischen Voraussetzungen eingegangen werden.

Ausgangspunkt der nachfolgenden Betrachtungen sei das folgende lineare Regressionsmodell

$$y_t = x_t'\beta_t + u_t \qquad t = 1, \ldots, T.$$

Dabei ist y_t der beobachtete Wert der abhaengigen Variablen zum Zeitpunkt t, x_t ein Kx1 Spaltenvektor mit den (zunaechst nichtstochastischen) Werten der K unabhaengigen Variablen (einschließlich Absolutglied), β_t ein (unbekannter) Kx1 Parametervektor und u_t eine normalverteilte Stoergroeße mit $E(u_t) = 0$ und Var $(u_t) = E(u_t^2) = \sigma^2$.

Die u_t, t=1, ..., T werden als unabhaengig vorausgesetzt.

Die zu testende Nullhypothese lautet nun:

H_0: $\beta_1 = \beta_2 = ... = \beta_T = \beta$ (Konstanz der Regressionskoeffizienten)

Damit laeßt sich das obige Modell in der ueblichen Form

$$y = X\beta + u$$

mit $y = (y_1, ..., y_T)'$, $X' = (x_1, ..., x_T)$ und $u = (u_1, ..., u_T)'$ schreiben.

Setzt man nun $X_r' = (x_1, ..., x_r)$ und $Y_r = (y_1, ..., y_r)'$, so lassen sich unter Einfuehrung der schrittweisen KQ-Schaetzer b_r, d.h.

$$b_r = (X_r' X_r)^{-1} X_r' Y_r \qquad r = K, K+1, ..., T$$

(dabei werde angenommen, daß rg $(X_r) = K$ ist, fuer r = K, ..., T), die "rekursiven Residuen" w_r definieren:

$$w_r = \frac{y_r - x_r' b_{r-1}}{(1 + x_r'(X_{r-1}' X_{r-1})^{-1} x_r)^{\frac{1}{2}}} \qquad \text{fuer } r=K+1, ..., T.$$

Bekanntlich handelt es sich bei den w_r um unabhaengige, normalverteilte Zufallsvariablen mit $E(w_r) = 0$ und $Var(w_r) = \sigma^2$.

Mit Hilfe der rekursiven Residuen lassen sich dann die CUSUM-Groessen W_r definieren

$$W_r = \frac{1}{\hat{\sigma}} \sum_{j=K+1}^{T} w_j \qquad r = K+1, \ldots, T$$

wobei $$\hat{\sigma} = \left(\frac{(y-X\hat{\beta})'(y-X\hat{\beta})}{T-K} \right)^{\frac{1}{2}}$$ ist

mit $\hat{\beta}$ als dem gewoehnlichen KQ-Schaetzer.

Zu einer vorgegebenen Irrtumswahrscheinlichkeit α koennen nun Parameter c und d angegeben werden, sodaß die Hypothese Ho genau dann abgelehnt wird, falls

$$\max_{K+1 \leq r \leq T} \left| W_r - (d + c(r-K)) \right| > 0 \quad \text{ist.}$$

(Naeheres dazu siehe Brown, Durbin, Evans (1975))

Streng genommen ist diese Vorgangsweise nur bei genuegend großem Stichprobenumfang T gerechtfertigt, da es sich beim CUSUM-Test um einen asymptotischen Test handelt, der lediglich die Bedingung

$$\lim_{T \to \infty} P\left(\max_{K+1 \leq r \leq T} \left| W_r - (d+c(r-K)) \right| > 0 \right) = \alpha \qquad \text{(gegeben Ho)}$$

erfuellt.

Nun zur Anwendung des CUSUM-Tests auf dynamische Modelle.

Ausgangspunkt fuer die Monte Carlo Experimente war ein einfaches dynamisches Regressionsmodell von der Form

$$y_t = ay_{t-1} + bx_t + c + u_t \qquad t = 1, \ldots, T \qquad (1)$$

Folgende zwei Testvarianten wurden dabei untersucht:

a) Nach einem Vorschlag von Dufour (1982, im weiteren auch Dufour-Test genannt) soll der CUSUM-Test auf die modifizierte Gleichung

$$y_t^* = bx_t + c + u_t^*,$$

$$\text{mit} \quad y_t^* = y_t - \hat{a}y_{t-1}$$

(wobei $\hat{a}$ fuer die KQ-Schaetzung von a steht) angewandt werden.

b) Der CUSUM-Test wird unmittelbar auf Gleichung (1) angewandt, also mit y_{t-1} als zusaetzlichem (jetzt stochastischen) Regressor.
(Diese Variante wird im folgenden als dynamischer CUSUM-Test bezeichnet)

Naehere Angaben zum "Design" der Monte Carlo Experimente finden sich in den beiden naechsten Abschnitten.

2. Empirische Signifikanzniveaus der beiden Testvarianten.

2.1. Der Dufour-Test.

Bei der Berechnung der empirischen Ablehnwahrscheinlichkeiten unter H_0 wurden verschiedene Stichprobenumfaenge T und verschiedene Regressionskoeffizienten a, b, c, beruecksichtigt. Der x-Regressor wurde mit $x_t = (-1)^t$, $t = 1, \ldots, T$, festgesetzt. Außerdem galt $y_0 = 0$. Fuer die Stoergroeßen u_t wurden unabhaengige $N(0,1)$-verteilte Zufallszahlen erzeugt. Die Anzahl der Simulationen fuer die einzelnen Kombinationen von T bzw. a, b, c betrug jeweils 1000. Fuer die Monte Carlo Experimente, deren Ergebnisse in den nachfolgenden Tabellen enthalten sind, wurden die ueblichen nominalen Signifikanzniveaus verwendet (α = 1 %, 5 %, 10 %).

Tabelle 1: Empirische Signifikanzniveaus fuer verschiedene Werte von T (a = 0.5, b = 1.0, c = 1.0)

α (in Prozent)	T = 30	60	120	1000
1	0.1	0.3	0.5	1.0
5	1.4	2.2	2.5	4.6
10	3.5	5.4	7.4	9.5

Tabelle 2: Empirische Signifikanzniveaus fuer verschiedene Werte von T bzw. a, b, c.

	T = 120 α (in Prozent) 1	5	10	T = 1000 α (in Prozent) 1	5	10
a = 0.5 b = 0.0 c = 0.0	0.5	2.3	7.9	0.6	3.8	8.8
a = 0.5 b = 2.0 c = 10.0	0.1	0.8	2.1	0.2	3.3	7.5

Tabelle 3: Empirische Signifikanzniveaus fuer verschiedene Werte von a (b = 1.0, c = 1.0, T = 120)

a	α (in Prozent) 1	5	10
-0.95	0.5	3.8	8.7
-0.90	0.7	3.8	8.4
-0.60	0.3	4.1	7.4
-0.30	0.7	3.1	8.2
0.00	0.5	3.9	7.7
0.30	0.7	3.3	6.9
0.60	0.4	2.6	6.3
0.90	0.3	2.0	3.5
0.95	0.1	1.0	1.5

Die Ergebnisse zeigen recht anschaulich, daß sich die empirischen Signifikanzniveaus bei zunehmenden Stichprobenumfang T immer mehr den vorgegebenen nominalen Signifikanzniveaus annaehern. Dabei spielen natuerlich auch der zugrundeliegende Zufallsmechanismus und wohl auch unterschiedliche Konvergenzgeschwindigkeiten (siehe z.B. Tabelle 2) eine gewisse Rolle. Die Resultate in Tabelle 3 weisen insbesondere auf den moeglichen Einfluß unterschiedlicher Werte fuer den Parameter a hin: Fuer $a \to 1$ scheint sich die Anpassung an die nominalen Signifikanzniveaus drastisch zu verschlechtern.

2.2 Der dynamische CUSUM-Test

Bei der Durchfuehrung des dynamischen CUSUM-Tests wurde zunaechst dasselbe "Design" wie beim Dufour-Test verwendet (siehe dazu die entsprechenden Angaben in 2.1.1.).

Dadurch besteht nun die Moeglichkeit, die Ergebnisse der folgenden Tabellen 4 - 6 mit denen der Tabellen 1 - 3 zu vergleichen.

Tabelle 4: Empirische Signifikanzniveaus fuer verschiedene Werte von T (a = 0.5, b = 1.0, c = 1.0)

α	T =			
(in Prozent)	30	60	120	1000
1	0.3	0.3	0.4	1.1
5	2.8	4.4	3.0	5.8
10	6.0	8.6	7.5	9.7

Tabelle 5: Empirische Signifikanzniveaus fuer verschiedene Werte von T bzw. a, b, c.

	T = 120 α (in Prozent)			T = 1000 α (in Prozent)		
	1	5	10	1	5	10
a = 0.5 b = 0.0 c = 0.0	0.8	5.5	9.3	0.9	7.1	12.2
a = 0.5 b = 2.0 c = 10.0	0.6	4.6	9.1	0.7	5.3	9.3

Tabelle 6: Empirische Signifikanzniveaus bei verschiedenen Werten des Parameters a ($T = 120$, $b = 1.0$, $c = 1.0$)

	α (in Prozent)		
a	1	5	10
-0.95	0.6	2.9	6.5
-0.90	0.4	3.3	6.8
-0.60	0.3	3.3	7.1
-0.30	0.3	2.8	5.8
0.0	0.3	3.1	7.6
0.30	0.5	3.4	7.3
0.60	0.4	4.3	9.6
0.90	0.9	5.2	10.1
0.95	0.8	4.5	8.4

Die Resultate der vorangegangenen Tabellen weisen fuer den dynamischen CUSUM-Test ein aehnliches Verhalten auf wie im Falle des Dufour-Tests, allerdings mit zwei bemerkenswerten Ausnahmen:

1) die Annaeherung an die nominalen Signifikanzniveaus erfolgt i.a. wesentlich rascher als beim Dufour-Test

2) in einigen Faellen werden diese Signifikanzniveaus sogar leicht ueberschritten.

Dieser letzte Punkt tritt besonders augenfaellig in der folgenden Tabelle zutage, wobei fuer den x-Regressor diesmal eine Trendvariable, d.h. $x_t = t$, verwendet wurde.

Tabelle 7: Empirische Signifikanzniveaus fuer verschiedene Werte von T bzw. a, b, c.

	T = 120 α (in Prozent)			T = 1000 α (in Prozent)		
	1	5	10	1	5	10
a = 0.5, b = 0.0, c = 0.0	1.0	5.1	10.7	1.1	5.6	10.1
a = 0.5, b = 1.0, c = 1.0	1.0	5.7	9.6	1.3	5.1	10.9
a = 0.5, b = 2.0, c = 10.0	0.4	3.8	7.6	0.7	4.2	10.8

Zusammenfassend kann gesagt werden:
Bei beiden Testvarianten zeigen die Ergebnisse der Monte Carlo Experimente, daß sich die empirischen Signifikanzniveaus bei zunehmendem Stichprobenumfang T immer besser bei den vorgegebenen Signifikanzniveaus "einpendeln".
(Der exakte Nachweis fuer dieses asymptotische Verhalten findet

sich in dem bereits erwaehnten Artikel von Kraemer, Ploberger, Alt (1985)).

Da die empirischen Resultate fuer den dynamischen CUSUM-Test eine hoehere Konvergenzgeschwindigkeit, d.h. eine schnellere Anpassung an das nominale Signifikanzniveau, ausweisen, sind wohl (fuer diesen Test) gewisse positive Auswirkungen auf die Teststaerke zu erwarten.

3. Empirische Untersuchungen zur Teststaerke.

Auf Grund der zahlreichen moeglichen Alternativen in Bezug auf Abweichungen von der Nullhypothese (Konstanz der Regressionskoeffizienten) war bei den "Power"-Studien eine Beschraenkung auf einige wenige, aber relevante Faelle unerlaeßlich. Die Festlegung der Alternativhypothesen erfolgte dabei unter folgenden Gesichtspunkten:

a) Fuer den gesamten Zeitbereich $t = 1, \ldots, T$ wird nur ein einziger Strukturbruch (structural shift) vorgegeben.

b) Der Zeitpunkt fuer das Auftreten dieses Strukturbruchs wird variabel gehalten.

c) Unterschiedliche Intensitaet und eine gewisse "Aufteilung" des Strukturbruchs auf verschiedene Parameter werden beruecksichtigt.

Es folgt eine genaue Beschreibung der Modellspezifikationen und der Alternativhypothesen.

Ausgangspunkt fuer die Monte Carlo Simulationen war wiederum das Modell

$$y_t = a\, y_{t-1} + bx_t + c + u_t \qquad t = 1, \ldots, T. \quad (1)$$

Bei den einzelnen Experimenten wurden verschiedene Stichprobenumfaenge ($T = 30, 60, 120$) und verschiedene nominale Signifikanzniveaus ($\alpha = 1\%, 5\%, 10\%$) verwendet. Groeβere Stichprobenumfaenge wurden nicht beruecksichtigt, da dies einen unverhaeltnismaeβigen Aufwand an Rechenzeit bedeutet haette.

Der x-Regressor wurde zunaechst mit $x_t = (-1)^t$ festgesetzt. Spaeter wurden fuer einige Experimente mit dem dynamischen CUSUM-Test $N(0,1)$-Zufallszahlen erzeugt, die dann im weiteren Verlauf der Untersuchungen konstant blieben.
Der Zeitpunkt T^* fuer den Strukturbruch konnte wie folgt variiert werden:

$$T^* = d \cdot T \qquad \text{mit } d = 0.3,\ 0.5,\ 0.7,$$

d.h. der Strukturbruch erfolgte in der ersten Haelfte, der Mitte oder der zweiten Haelfte des "Beobachtungszeitraums" $t = 1, \ldots, T$.

Nun zur eigentlichen Konstruktion des Strukturbruchs.
Fuer die Zeitpunkte $t = 1, \ldots, T^*$ wurde die obige Modellgleichung unveraendert uebernommen, d.h. es galt

$$y_t = ay_{t-1} + bx_t + c + u_t \quad \text{fuer } t = 1, \ldots T^* .$$

(Als Stoergroeβen u_t wurden wieder unabhaengige $N(0,1)$-Zufallszahlen verwendet. Die Parameter a, b, und c wurden aehnlich wie

in Abschnitt 1 variiert. Genauere Angaben dazu finden sich in den Ueberschriften der nachfolgenden Tabellen.)
Ab dem Zeitpunkt $T^* + 1$ galt dann

$$y_t = ay_{t-1} + \left(b + \frac{b}{\sqrt{T}} \sin \psi\right) x_t + \left(c + \frac{b}{\sqrt{T}} \cos \psi\right) + u_t$$

wobei ψ und b^* folgende Werte annehmen koennen:

$$\psi = 0^\circ,\ 36^\circ,\ 90^\circ$$
$$b^* = 4.8,\ 7.2,\ 9.6,\ 12.0 \quad .$$

Bei gegebenen Werten fuer T, a, b und c konnten somit durch entsprechende Variation der Groeßen d, b^* und ψ insgesamt 36 verschiedene Kombinationsmoeglichkeiten untersucht werden. Dabei wurden jeweils 1000 Simulationen durchgefuehrt.
Aus Platzgruenden und aus Gruenden einer uebersichtlichen Darstellung sind in den folgenden Tabellen ausschließlich Simulationsergebnisse fuer ein nominales Signifikanzniveau $\alpha = 5\%$ enthalten. Außerdem wurden nur bei den Tabellen 8 - 10 unterschiedliche Stichprobenumfaenge beruecksichtigt (auf Anfrage koennen die entsprechenden Ergebnisse fuer $\alpha = 1\%$, bzw. 10% und $T = 30$ bzw. 60 vom Autor bezogen werden).

Tabelle 8: Empirische Teststaerke fuer T = 30 *)

($x_t = (-1)^t$ / a= 0.5, b = 1.0, c = 1.0)

d	b^*	Dufour-Test Ψ 0	36	90	Dynamischer CUSUM-Test Ψ 0	36	90
0.3	4.8	0.3	0.3	1.4	2.3	1.6	1.5
	7.2	0.1	0.1	2.0	3.3	3.1	1.8
	9.6	0.0	0.1	2.8	2.4	3.7	1.4
	12.0	0.0	0.0	4.8	2.0	3.5	2.7
0.5	4.8	0.3	0.3	0.6	2.1	1.4	1.6
	7.2	0.3	0.2	1.4	2.8	2.2	0.7
	9.6	0.1	0.0	3.4	3.1	1.4	0.2
	12.0	0.2	0.0	5.5	2.0	1.5	0.2
0.7	4.8	0.7	0.7	1.1	2.5	1.6	1.7
	7.2	0.3	0.4	0.8	2.3	1.1	1.2
	9.6	0.4	0.4	1.5	1.8	0.5	0.5
	12.0	0.3	0.0	3.1	2.6	0.6	0.0

*) Angaben in Prozent

Tabelle 9: Empirische Teststaerke fuer T = 60 *)

($x_t = (-1)^t$ / a = 0.5, b = 1.0, c = 1.0)

d	b^*	Dufour-Test ψ			Dynamischer CUSUM-Test ψ		
		0	36	90	0	36	90
0.3	4.8	0.7	0.8	2.2	6.3	5.2	3.4
	7.2	0.6	1.1	2.7	10.4	9.3	2.7
	9.6	0.1	0.5	4.2	11.1	11.8	2.6
	12.0	0.1	0.6	6.0	9.6	12.8	3.2
0.5	4.8	0.9	2.0	1.9	6.3	4.4	2.6
	7.2	0.4	0.5	2.2	9.2	6.3	1.6
	9.6	0.3	0.2	4.1	11.5	9.7	1.5
	12.0	0.1	0.4	6.5	9.3	11.7	1.2
0.7	4.8	1.0	1.0	2.5	3.7	3.2	2.5
	7.2	0.4	0.9	3.1	5.7	2.4	1.9
	9.6	0.5	0.1	2.9	7.2	4.7	1.3
	12.0	0.3	0.1	4.1	6.7	3.1	1.2

*) Angaben in Prozent

Tabelle 10: Empirische Teststaerke fuer T = 120 *)

($x_t = (-1)^t$ / a = 0.5, b = 1.0, c = 1.0)

d	b^*	Dufour-Test ψ 0	36	90	Dynamischer CUSUM-Test ψ 0	36	90
0.3	4.8	6.9	6.1	3.4	10.5	7.7	5.3
	7.2	10.3	10.5	3.0	25.2	17.9	3.3
	9.6	8.8	15.6	3.7	32.8	28.0	5.2
	12.0	8.9	18.7	5.4	37.7	37.8	3.8
0.5	4.8	5.7	2.8	2.9	11.5	8.7	3.5
	7.2	3.1	4.6	4.1	19.5	13.4	2.4
	9.6	3.7	5.8	4.8	30.1	22.4	3.5
	12.0	1.9	6.3	4.2	33.5	32.6	1.4
0.7	4.8	2.1	2.4	4.1	6.2	4.8	3.1
	7.2	2.0	1.8	3.1	8.2	6.6	2.6
	9.6	1.0	1.3	4.3	11.5	9.6	2.3
	12.0	0.3	1.2	4.3	19.8	12.7	1.1

*) Angaben in Prozent

Den voranstehenden Tabellen lassen sich unschwer folgende Informationen entnehmen:

a) Fuer den Fall $\psi = 0°$ bzw. $\psi = 36°$ weisen die Ergebnisse fuer den dynamischen CUSUM-Test gegenueber dem Dufour-Test eine ungleich groeßere Teststaerke auf. Mit zunehmenden Stichprobenumfang T zeigen sich z.T. betraechtliche Unterschiede.

b) Ist $\psi = 90°$, so laeßt sich ein leichtes "Uebergewicht" des Dufour-Tests feststellen. Abgesehen davon, daß die Unterschiede nicht dramatisch ausfallen, zeigt sich hier bei beiden Testvarianten, daß die Teststaerke i.a. unterhalb des vorgegebenen Signifikanzniveaus liegt.
(Wegen $\frac{b^*}{\sqrt{T}} \cos 90° = 0$ tritt uebrigens der Strukturbruch ausschließlich beim Koeffizienten des x-Regressors auf).

c) Mit zunehmendem b^* sollte sinnvollerweise auch die Teststaerke zunehmen, da der Strukturbruch dann umso deutlicher ausfaellt. Abgesehen vom Fall $\psi = 90°$ zeigt sich beim Dufour-Test recht haeufig das "paradoxe" Resultat , daß sich die Teststaerke dann mehr und mehr verringert.

d) Die Ergebnisse beider Testvarianten zeigen außerdem, daß sich bei einem spaeten Auftreten des Strukturbruchs (d = 0.7) die Chancen fuer seine "Entdeckung" wesentlich verschlechtern.

In der Folge wurden fuer den dynamischen CUSUM-Test noch einige weitere Simulationsexperimente durchgefuehrt. Stellvertretend dafuer seien die Ergebnisse der umstehenden Tabelle genannt. Dabei wurden fuer den x-Regressor zunaechst N(0,1)-Zufallszahlen erzeugt, die dann bei saemtlichen Simulationen konstant blieben (Somit galt x_t als nichtstochastische Variable). Die Parameterkonstellationen wurden wie folgt gewaehlt:

$$a = 0.5,\ b = 1.0,\ c = 1.0$$

$$\text{bzw. } a = 0.5,\ b = 2.0,\ c = 10.0$$

Der Stichprobenumfang wurde mit $T = 120$, das Signifikanzniveau mit $\alpha = 5\%$ festgesetzt. Das Konzept zur Konstruktion des Strukturbruchs wurde von den frueheren Experimenten unveraendert uebernommen. Die Anzahl der Simulationen pro Kombinationsmoeglichkeit betrug wiederum jeweils 1000.

Tabelle 11: Empirische Teststaerke fuer den dynamischen CUSUM-Test (X_t: N(0,1) / T = 120, a = 0.5) *)

		b = 1.0, c = 1.0			b = 2.0, c = 10.0		
		ψ			ψ		
d	b^*	0	36	90	0	36	90
0.3	4.8	16.5	13.0	2.9	20.7	11.5	3.6
	7.2	35.0	23.5	3.0	42.6	19.6	3.3
	9.6	49.4	33.0	2.0	63.8	37.6	3.7
	12.0	60.0	45.0	1.7	83.2	51.9	2.6
0.5	4.8	13.3	9.1	4.4	15.8	8.2	4.3
	7.2	24.2	16.2	3.5	28.9	17.6	2.1
	9.6	37.4	27.3	2.2	55.0	28.6	3.1
	12.0	50.4	36.0	2.4	70.8	45.4	2.0
0.7	4.8	6.9	6.2	2.8	7.2	5.5	3.9
	7.2	13.2	9.8	3.0	12.0	8.2	2.9
	9.6	19.7	14.1	4.4	20.4	13.4	3.1
	12.0	27.2	21.4	2.2	36.0	19.0	1.9

*) Angaben in Prozent

Die Ergebnisse zeigen aehnliche Regelmaeßigkeiten wie bei den frueheren Experimenten. So liegt im Falle $\psi = 90°$ die Teststaerke generell unterhalb des Signifikanzniveaus. Fuer $\psi = 0°$ bzw. $\psi = 36°$ nimmt mit zunehmendem b^* die Teststaerke ebenfalls zu. Allerdings verschlechtern sich die Ergebnisse sehr rasch bei einem spaeteren Auftreten des Strukturbruchs.
Eine genauere Untersuchung dieses Phaenomens waere sicherlich von großem Interesse, da Strukturbrueche, die erst gegen Ende des Beobachtungszeitraumes auftreten, die zukuenftige Entwicklung besonders nachhaltig beeinflussen koennen.

II. Eine Modifikation des CUSUM-Tests.

1. Einfuehrung einer neuen Varianzschaetzung.

"It seems much more sensible to me, however, to estimate σ by the square root of $\sum (wj - \overline{w})^2 / (T-k-1)$. This does not affect the theory behind the cusum test, but it is likely to make the procedure more effective as the cusum will tend to be larger in absolute value under the alternative hypothesis."

Diese sicherlich interessante Bemerkung wurde von Harvey im Diskussionsteil des Artikels von Brown, Durbin, Evans gemacht (S. 179; dabei bedeutet $\overline{w}$ das arithmetische Mittel der rekursiven Residuen und k die Anzahl der Regressoren. Siehe dazu auch die entsprechenden Definitionen in Kapitel I, Abschnitt 1). Allerdings scheint Harvey's Vorschlag seit Erscheinen des obigen Artikels keine allzu große Beachtung gefunden zu haben.

Im Folgenden soll nun durch eine Reihe von Monte Carlo Experimenten der Einfluß der modifizierten Varianzschaetzung auf den CUSUM-Test genauer untersucht werden.

Um einen unmittelbaren Vergleich zwischen dem "gewoehnlichen" und dem "modifizierten" CUSUM-Test zu ermoeglichen, wurden bei jeder einzelnen Simulation die zufaellig erzeugten y-Werte beiden Testvarianten "gleichzeitig" unterworfen, d.h. fuer beide Testvarianten wurden jeweils derselbe Datensatz verwendet.

(Auf Grund eines Mißverstaendnisses des Autors wurde fuer den Nenner der neuen Varianzschaetzung T-K statt T-K-1 verwendet. Dies duerfte den aufschlußreichen Ergebnissen der beiden naechsten Abschnitte aber keinen nennenswerten Abbruch tun.)

2. Vergleich der empirischen Signifikanzniveaus.

Zum Vergleich der Signifikanzniveaus wurde das folgende (statische) Modell herangezogen:

$$y_t = bx_t + c + u_t \qquad t = 1, \ldots, T.$$

Als Stichprobenumfang wurde T = 10, 20, 30, 60, 120, 1000 verwendet, der x-Regressor mit $x_t = (-1)^t$, die Parameter mit b = 1.0 und c = 1.0 festgesetzt. Die Anzahl der Simulationen fuer jedes T betrug 1000. Fuer die Stoergroeßen wurden wieder N(0,1)-Zufallszahlen erzeugt. Außerdem wurden die ueblichen nominalen Signifikanzniveaus verwendet (α = 1 % , 5 % , 10 %).

Tabelle 12: Empirische Signifikanzniveaus fuer verschiedene Werte von T.

T	CUSUM-Test α (in Prozent) 1	5	10	Modifizierter CUSUM-Test α (in Prozent) 1	5	10
10	0.0	0.0	1.3	2.1	5.5	9.2
20	0.0	1.6	4.0	0.7	3.7	5.9
30	0.1	2.0	6.1	0.6	4.7	8.1
60	0.5	4.2	8.4	1.0	4.9	9.6
120	0.3	3.7	8.1	0.6	3.9	8.4
1000	1.4	5.8	11.0	1.4	5.8	11.0

Es faellt sofort auf, daß bei "kleinen" Stichprobenumfaengen, wie sie insbesondere bei oekonomischen Zeitreihen vorkommen, die Anpassung des modifizierten CUSUM-Tests an die nominalen Signifikanzniveaus z.T. erheblich besser ist als beim gewoehnlichen CUSUM-Test. Dies duerfte nicht ohne Einfluß auf die Teststaerke sein, wie die nachfolgenden "Power"-Studien noch zeigen werden. Da die beiden Tests dasselbe asymptotische Verhalten besitzen, ist es nicht ueberraschend, daß sich die Ergebnisse bei großem T nur noch wenig voneinander unterscheiden.

3. Vergleich der empirischen Teststaerken.

Ein erster Vergleich der Teststaerken basierte wiederum auf dem in Kapitel I, Abschnitt 2, eingefuehrten Konzept eines einzigen Strukturbruchs. Folgende Stichprobenumfaenge wurden verwendet:

$$T = 10,\ 20,\ 30,\ 60 \text{ und } 120\,.$$

Ansonsten galten dieselben Voraussetzungen wie im letzten Abschnitt, d.h. $x_t = (-1)^t$, $b = 1.0$ und $c = 1.0$. Fuer die Stoergroeßen wurden wieder $N(0,1)$-Zufallszahlen erzeugt.

Auf den folgenden Seiten sind die Ergebnisse der Simulationsexperimente in uebersichtlicher Form zusammengestellt.

Tabelle 13: Empirische Teststaerke fuer T = 10 *)

($x_t = (-1)^t$ / b = 1.0, c = 1.0)

d	b^*	CUSUM-Test Ψ 0	36	90	Modifizierter CUSUM-Test Ψ 0	36	90
0.3	4.8	0.2	0.1	0.0	34.9	22.3	2.3
	7.2	0.3	0.2	0.0	57.0	29.8	0.4
	9.6	0.0	0.1	0.0	71.3	40.5	0.1
	12.0	0.2	0.1	0.0	87.4	48.4	0.0
0.5	4.8	0.0	0.0	0.0	19.9	14.9	1.8
	7.2	0.0	0.0	0.0	32.7	17.7	0.2
	9.6	0.0	0.0	0.0	42.6	20.3	0.2
	12.0	0.0	0.0	0.0	50.6	18.5	0.0
0.7	4.8	0.0	0.0	0.0	9.9	6.8	2.1
	7.2	0.0	0.0	0.0	11.0	6.9	1.1
	9.6	0.0	0.0	0.0	8.9	5.2	0.2
	12.0	0.0	0.0	0.0	10.6	3.4	0.0

*) Angaben in Prozent

Tabelle 14: Empirische Teststaerke fuer T = 20 *)
($x_t = (-1)^t$ / b = 1.0, c = 1.0)

d	b^*	CUSUM-Test ψ 0	36	90	Modifizierter CUSUM-Test ψ 0	36	90
0.3	4.8	12.1	4.6	1.1	27.7	13.9	3.1
	7.2	27.4	11.0	0.6	52.3	29.4	1.1
	9.6	44.9	15.5	0.2	75.8	42.3	0.5
	12.0	61.5	22.6	0.1	89.5	55.6	0.2
0.5	4.8	5.4	2.6	0.2	18.1	9.7	1.5
	7.2	12.0	3.9	0.5	34.0	16.1	1.2
	9.6	19.7	4.1	0.1	54.4	23.5	0.3
	12.0	28.9	4.7	0.0	70.3	29.3	0.0
0.7	4.8	1.8	1.4	0.7	8.2	5.7	2.2
	7.2	1.9	1.0	0.2	9.6	6.3	1.1
	9.6	3.7	0.4	0.3	20.1	6.7	0.4
	12.0	3.2	0.7	0.0	21.8	6.8	0.0

*) Angaben in Prozent

Tabelle 15: Empirische Teststaerke fuer T = 30 *)
($x_t = (-1)^t$ / b = 1.0, c = 1.0)

d	b^*	CUSUM-Test Ψ 0	36	90	Modifizierter CUSUM-Test Ψ 0	36	90
0.3	4.8	16.1	10.6	1.4	26.2	18.1	2.6
	7.2	40.8	19.8	0.6	55.9	33.1	1.4
	9.6	62.3	29.7	0.2	78.4	48.2	0.4
	12.0	81.1	48.0	0.0	92.2	67.5	0.1
0.5	4.8	10.1	5.6	1.2	17.2	12.3	2.5
	7.2	22.0	9.4	0.6	35.7	19.5	1.2
	9.6	38.5	17.5	0.2	60.3	32.5	0.8
	12.0	55.4	25.0	0.0	76.4	44.6	0.1
0.7	4.8	3.2	3.1	1.0	6.7	6.6	2.8
	7.2	4.2	3.0	0.6	10.9	6.7	1.5
	9.6	9.4	3.6	0.2	21.2	11.6	0.5
	12.0	14.0	4.5	0.0	31.9	16.2	0.3

*) Angaben in Prozent

Tabelle 16: Empirische Teststaerke fuer T = 60 *)

($x_t = (-1)^t$ / b = 1.0, c = 1.0)

d	b^*	CUSUM-Test Ψ 0	36	90	Modifizierter CUSUM-Test Ψ 0	36	90
0.3	4.8	21.7	12.6	2.6	25.9	16.0	3.3
	7.2	47.4	31.4	1.5	54.1	36.7	1.8
	9.6	73.1	47.2	1.0	78.7	55.8	1.2
	12.0	92.1	65.6	0.8	94.3	72.4	0.9
0.5	4.8	14.4	8.6	2.2	19.3	11.4	3.1
	7.2	32.0	16.9	1.3	37.1	21.3	1.8
	9.6	56.8	32.1	0.7	65.3	39.4	1.0
	12.0	78.4	45.0	0.4	85.8	54.4	0.5
0.7	4.8	4.4	4.9	2.2	5.9	6.3	3.1
	7.2	9.5	7.7	1.5	13.7	10.7	1.8
	9.6	19.0	10.5	1.1	26.5	15.3	1.2
	12.0	31.2	13.8	0.3	39.7	19.1	0.7

*) Angaben in Prozent

Tabelle 17: Empirische Teststaerke fuer T = 120 *)

($x_t = (-1)^t$ / b = 1.0, c = 1.0)

d	b^*	CUSUM-Test Ψ 0	36	90	Modifizierter CUSUM-Test Ψ 0	36	90
0.3	4.8	22.2	15.4	2.5	23.9	16.7	3.0
	7.2	53.5	32.8	2.8	57.0	35.9	3.0
	9.6	80.3	56.1	2.0	82.7	59.6	2.2
	12.0	94.0	78.6	0.8	95.0	81.1	1.1
0.5	4.8	15.8	8.9	2.7	17.8	9.7	3.2
	7.2	40.6	20.6	2.4	43.4	24.0	2.6
	9.6	65.5	41.7	2.3	69.1	45.1	2.6
	12.0	88.5	61.1	1.5	91.1	66.9	1.9
0.7	4.8	7.0	5.1	3.3	8.2	5.7	3.4
	7.2	15.1	9.7	2.4	17.5	11.6	2.6
	9.6	28.8	12.9	1.9	33.6	16.0	2.1
	12.0	47.8	25.4	1.2	51.9	29.0	1.4

*) Angaben in Prozent

Die Ergebnisse zeigen eine generelle Ueberlegenheit des modifizierten CUSUM-Tests, die insbesondere bei kleinen Stichprobenumfaengen wie $T = 10$ bzw. $T = 20$ besonders auffaellig ist.

Fuer eine Folge von weiteren Experimenten wurde eine neue Variante eingefuehrt. Dabei treten nun waehrend des gesamten Beobachtungszeitraumes genau zwei Strukturbrueche auf und zwar zu den Zeitpunkten

$$T_1^* = 0.3 \cdot T$$
$$\text{und} \quad T_2^* = 0.7 \cdot T$$

Das Modell wurde dann in seinem Zeitverlauf folgendermaßen festgesetzt:

<u>Variante A</u>

$$y_t = bx_t + c + u_t \qquad \text{fuer } t = 1,\dots,T_1^*$$
$$y_t = (b + \frac{b^*}{\sqrt{T}} \sin\psi)x_t + (c + \frac{b^*}{\sqrt{T}} \cos\psi) + u_t \quad \text{fuer } t = T_1^* + 1,\dots T_2^*$$
$$y_t = (b + 2\frac{b^*}{\sqrt{T}} \sin\psi)x_t + (c + 2\frac{b^*}{\sqrt{T}} \cos\psi) + u_t \quad \text{fuer } T_2^*+1,\dots,T.$$

Die Definitionen der einzelnen Parameter und Variablen wurden von den letzten Experimenten unveraendert uebernommen und auch weiterhin beibehalten.

Tabelle 18: Empirische Teststaerke fuer T = 10 *)
(Variante A)

	CUSUM-Test ψ			Modifizierter CUSUM-Test ψ		
b^*	0	36	90	0	36	90
4.8	0.0	0.0	0.0	58.0	30.3	0.5
7.2	0.0	0.0	0.0	79.7	41.6	0.0
9.6	0.0	0.0	0.0	95.1	50.4	0.0
12.0	0.0	0.0	0.0	98.4	57.0	0.0

Tabelle 19: Empirische Teststaerke fuer T = 20 *)
(Variante A)

	CUSUM-Test ψ			Modifizierter CUSUM-Test ψ		
b^*	0	36	90	0	36	90
4.8	22.8	6.5	0.1	56.4	26.7	1.0
7.2	49.4	13.6	0.0	84.5	48.7	0.1
9.6	78.4	18.7	0.0	97.4	65.0	0.0
12.0	92.0	22.9	0.0	99.9	78.2	0.0

*) Angaben in Prozent

Tabelle 20: Empirische Teststaerke fuer T = 30 *)

(Variante A)

b^*	CUSUM-Test Ψ 0	36	90	Modifizierter CUSUM-Test Ψ 0	36	90
4.8	37.3	16.1	0.4	55.6	30.2	0.8
7.2	75.0	36.0	0.1	90.1	60.3	0.1
9.6	94.8	57.8	0.0	98.4	80.7	0.1
12.0	99.3	74.0	0.0	100.0	91.8	0.0

Tabelle 21: Empirische Teststaerke fuer T = 60 *)

(Variante A)

b^*	CUSUM-Test Ψ 0	36	90	Modifizierter CUSUM-Test Ψ 0	36	90
4.8	48.6	27.1	1.6	56.6	34.1	2.5
7.2	88.6	57.1	0.4	93.0	66.8	0.8
9.6	98.6	85.0	0.0	99.5	89.7	0.1
12.0	100.0	96.5	0.0	100.0	98.5	0.0

*) Angaben in Prozent

Tabelle 22: Empirische Teststaerke fuer T = 120 *)

(Variante A)

	CUSUM-Test ψ			Modifizierter CUSUM-Test ψ		
b^*	0	36	90	0	36	90
4.8	55.1	32.5	2.6	57.9	36.9	3.2
7.2	91.2	71.7	1.9	93.0	75.0	2.0
9.6	99.8	92.6	0.8	99.8	94.1	0.8
12.0	100.0	99.6	0.1	100.0	99.7	0.1

*) Angaben in Prozent

Wie schon zuvor zeigt die Teststaerke des modifizierten CUSUM-Tests ein deutliches Uebergewicht bei kleinen Stichproben. Außerdem sind weiterhin einige Regelmaeßigkeiten zu beobachten, auf die schon an anderer Stelle hingewiesen wurde, so z.B. eine aeußerst geringe Teststaerke bei $\psi = 90°$.

Fuer die naechsten Experimente wurde die Richtung des Strukturbruchs geaendert:

Variante B

$$y_t = bx_t + c + u_t \quad \text{fuer } t = 1,\dots,T_1^*$$

$$y_t = \left(b + \frac{b^*}{\sqrt{T}}\sin\psi\right)x_t + \left(c + \frac{b^*}{\sqrt{T}}\cos\psi\right) + u_t \quad \text{fuer } t = T_1^*+1,\dots,T_2^*$$

$$y_t = \left(b - \frac{b^*}{\sqrt{T}}\sin\psi\right)x_t + \left(c - \frac{b^*}{\sqrt{T}}\cos\psi\right) + u_t \quad \text{fuer } t = T_2^*+1,\dots,T.$$

Tabelle 23: Empirische Teststaerke fuer T = 10 *)
(Variante B)

b^*	CUSUM-Test Ψ 0	36	90	Modifizierter CUSUM-Test Ψ 0	36	90
4.8	0.0	0.0	0.0	0.0	0.3	0.6
7.2	0.0	0.0	0.0	0.0	0.0	0.0
9.6	0.0	0.0	0.0	0.0	0.0	0.0
12.0	0.0	0.0	0.0	0.0	0.0	0.0

Tabelle 24: Empirische Teststaerke fuer T = 20 *)
(Variante B)

b^*	CUSUM-Test Ψ 0	36	90	Modifizierter CUSUM-Test Ψ 0	36	90
4.8	0.3	0.4	0.3	0.7	1.0	0.6
7.2	0.3	0.1	0.0	0.4	0.3	0.1
9.6	0.1	0.0	0.0	0.1	0.1	0.0
12.0	0.0	0.0	0.0	0.0	0.0	0.0

*) Angaben in Prozent

Tabelle 25: Empirische Teststaerke fuer T = 30 *)
(Variante B)

b^*	CUSUM-Test ψ 0	36	90	Modifizierter CUSUM-Test ψ 0	36	90
4.8	3.5	2.8	0.3	4.3	3.5	1.2
7.2	3.8	1.8	0.0	4.2	2.0	0.0
9.6	3.4	0.7	0.0	3.5	0.8	0.0
12.0	2.2	0.4	0.0	2.2	0.4	0.0

Tabelle 26: Empirische Teststaerke fuer T = 60 *)
(Variante B)

b^*	CUSUM-Test ψ 0	36	90	Modifizierter CUSUM-Test ψ 0	36	90
4.8	9.3	7.2	1.3	10.0	7.8	1.6
7.2	13.8	8.1	0.2	14.1	8.5	0.2
9.6	19.3	8.2	0.1	20.2	8.5	0.1
12.0	25.4	7.2	0.0	25.8	7.2	0.1

*) Angaben in Prozent

Tabelle 27: Empirische Teststaerke fuer T = 120 *)
(Variante B)

b^*	CUSUM-Test ψ 0	36	90	Modifizierter CUSUM-Test ψ 0	36	90
4.8	13.6	9.3	2.1	14.0	9.9	2.5
7.2	28.3	15.9	1.3	29.0	16.7	1.4
9.6	42.2	22.2	0.8	43.5	22.7	0.9
12.0	55.6	30.6	0.2	57.8	31.3	0.3

*) Angaben in Prozent

Die Resultate fuer die Variante B zeigen, daß beide Tests nur sehr unzureichend auf die "entgegengesetzt verlaufenden" Strukturbrueche reagieren. Entscheidend aber bleibt weiterhin, daß der gewoehnliche CUSUM-Test gegenueber dem modifizierten Test (leicht) abfaellt.

Abschließend soll noch eine dritte Variante untersucht werden, bei der die Regressionskoeffizienten einen "Random walk" bilden:

Variante C

$y_t = bx_t + c + u_t$ fuer $t = 1,2$

$y_t = b_t x_t + c_t + u_t$ fuer $t = 3, \ldots, T$

wobei $\quad b_t = b_{t-1} + \frac{b^*}{\sqrt{T}} \sin\psi \cdot \varepsilon_t,$

$$c_t = c_{t-1} + \frac{b^*}{\sqrt{T}} \cos\psi \cdot \varepsilon_t \qquad \text{mit } \varepsilon_t \sim N(0,1)$$

und $\quad b_1 = b_2 = b \quad$ bzw. $\quad c_1 = c_2 = c$ ist.

<u>Tabelle 28:</u> Empirische Teststaerke fuer T = 10 *)

(Variante C)

	CUSUM-Test ψ			Modifizierter CUSUM-Test ψ		
b^*	0	36	90	0	36	90
4.8	0.2	0.0	0.0	40.4	17.7	0.4
7.2	0.2	0.0	0.0	44.1	18.9	0.0
9.6	0.0	0.0	0.0	46.1	19.3	0.0
12.0	0.6	0.0	0.0	45.8	18.6	0.1

*) Angaben in Prozent

Tabelle 29: Empirische Teststaerke fuer T = 20 *)

(Variante C)

b^*	CUSUM-Test ψ 0	36	90	Modifizierter CUSUM-Test ψ 0	36	90
4.8	40.9	16.5	0.1	53.5	35.8	0.2
7.2	47.2	18.3	0.0	60.1	39.9	0.0
9.6	49.8	20.5	0.0	62.1	42.3	0.0
12.0	49.7	24.7	0.0	63.4	47.5	0.0

Tabelle 30: Empirische Teststaerke fuer T = 30 *)

(Variante C)

b^*	CUSUM-Test ψ 0	36	90	Modifizierter CUSUM-Test ψ 0	36	90
4.8	56.5	38.3	0.0	62.1	48.3	0.2
7.2	64.3	47.1	0.0	70.2	55.4	0.0
9.6	64.1	49.4	0.0	70.3	58.6	0.0
12.0	68.2	49.8	0.0	73.7	58.5	0.0

*) Angaben in Prozent

Tabelle 31: Empirische Teststaerke fuer T = 60 *)

(Variante C)

b^*	CUSUM-Test ψ			Modifizierter CUSUM-Test ψ		
	0	36	90	0	36	90
4.8	80.9	67.6	0.0	83.4	70.4	0.0
7.2	85.5	74.3	0.0	88.0	77.1	0.0
9.6	88.6	73.9	0.0	90.1	76.7	0.0
12.0	88.1	76.4	0.0	89.3	78.9	0.0

Tabelle 32: Empirische Teststaerke fuer T = 120 *)

(Variante C)

b^*	CUSUM-Test ψ			Modifizierter CUSUM-Test ψ		
	0	36	90	0	36	90
4.8	94.7	87.5	0.2	95.1	88.4	0.2
7.2	97.0	91.7	0.1	97.3	92.2	0.1
9.6	97.3	92.6	0.0	97.4	93.1	0.0
12.0	97.6	91.5	0.0	97.9	92.3	0.0

*) Angaben in Prozent

Die Ergebnisse dieses Kapitels legen den Schluß nahe, daß der durch Einfuehrung einer neuen Varianzschaetzung modifizierte CUSUM-Test dem herkoemmlichen CUSUM-Test zumindest ebenbuertig ist. Auf Grund der z.T. eklatanten Unterschiede bei kleinen Stichproben, die gerade fuer die oekonomischen Anwendungen besonders typisch sind, sollte der modifizierte Test bevorzugt werden.

Literatur

Brown, Durbin, Evans, Techniques for Testing the Constancy of Regression Relationships over Time (with Discussion), Journal of the Royal Statistical Society B 37, 1975, 149 - 192.

Dufour, Recursive Stability Analysis of Linear Regression Relationships, Journal of Econometrics 19, 1982, 31 - 76.

Kraemer, Ploberger, Alt, Testing for Structural Change in Dynamic Models (unveroeffentlichtes Manuskript)