

Testing for Multivariate Cointegration in the Presence of Structural Breaks: p-Values and Critical Values

David E. Giles & Ryan T. Godwin

Department of Economics, University of Victoria

June, 2011

Abstract

Testing for multivariate cointegration when the data exhibit structural breaks is a problem that is encountered frequently in empirical economic analysis. The standard tests must be modified in this situation, and the asymptotic distributions of the test statistics change accordingly. We supply code that allows practitioners to easily calculate both p-values and critical values for the trace tests of Johansen *et al.* (2000). Access is also provided to tables of critical values for a broad selection of situations.

Keywords Cointegration, structural breaks, trace test, p-values, critical values

JEL Classifications C12, C32, C87

Contact author:

David E. Giles, Dept. of Economics, University of Victoria, P.O. Box 1700, STN CSC, Victoria, B.C., Canada V8W 2Y2; e-mail: dgiles@uvic.ca; Phone: (250) 721-8540; FAX: (250) 721-6214

1. Introduction

The need to test for unit roots and for cointegration in the presence of structural breaks arises frequently in practice when modelling time-series data. There is a substantial literature on unit root testing when structural breaks are present (*e.g.*, see the surveys by Perron, 2006; Byrne and Perman, 2007; Glynn *et al.*, 2007). However, fewer tools are available for multivariate cointegration testing, of the Johansen (1988) type, in the context of structural breaks. Noteworthy exceptions are the contributions of Inoue (1999) and Johansen *et al.* (2000). The latter authors adapt the basic Johansen framework to allow for trend and level breaks at several known points, and a very readable account of the application of the Johansen *et al.* methodology is provided by Joyeux (2007).

The asymptotic distributions of the cointegration tests developed by Johansen *et al.* are not known. Johansen *et al.* (2000, pp. 226-230) discuss the calculation of appropriate critical values on the basis of a simulation response surface, and approximation using a Gamma distribution. However, the implementation of their calculations is far from trivial. In addition, the information needed to compute p -values is not provided in the literature. So, in this note we provide transportable computer code for computing both p -values and critical values for these tests, and tabulate some selected critical values. In addition, links are provided for downloading more extensive tables of critical values.

2. Tests

Two variants of the usual trace test for cointegration among p time-series are considered by Johansen *et al.* (2000). These are the $H_l(r)$ and $H_c(r)$ tests for when there are $(q - 1)$ breaks (*i.e.*, q sub-samples) in a linear trend or in a constant level of the data, respectively. Here, r denotes the cointegrating rank. The asymptotic distributions of the test statistics depend on the values of $(p - r)$ and the locations of the break-points in the sample. These break-points are denoted $v_j = (T_j / T)$, where T is the full sample size and T_j is the last observation of the j^{th} sub-sample; $j = 1, 2, \dots, q$.

3. p-values and Critical values

We have written code for the R programming environment (R Development Core Team, 2011) and for the EViews econometric package (IHS EViews, 2011), that allows the user to generate asymptotic p -values and critical values for both the $H_l(r)$ and $H_c(r)$ tests, for user-selected significance levels, break-points, number of variables, and degree of cointegration. The programs use the estimated response surfaces in Table 4 of Johansen *et al.* (2000) to obtain the mean and variance of the asymptotic distributions. These two moments are then used to “recover” the shape and scale parameters of a Gamma distribution, which in turn is used to approximate the asymptotic sampling distributions

themselves. From these, both p -values and percentiles are computed. When there is a single break ($q = 2$), the percentiles are symmetric with respect to ν_1 . That is, the critical values for $(1 - \nu_1)$ are the same as those for ν_1 . When there are two structural breaks ($q = 3$), for a fixed value of ν_2 the critical values are symmetric about $(\nu_2/2)$ with respect to the value of ν_1 . For example, if $\nu_2 = 0.3$ then the percentiles for $\nu_1 = 0.05$ are the same as those for $\nu_1 = 0.25$, for both tests. In addition, when $q = 3$, for any choice of ν_1 and ν_2 , the critical values are the same if the break-point proportions are changed to ν_1^* and ν_2^* , where ν_1^* is any element of the *ordered* set $V_1 = \{\nu_1, (\nu_2 - \nu_1), (1 - \nu_2)\}$; and ν_2^* is any element of the ordered set $V_2 = \{(1 - \nu_1), (1 - \nu_2 + \nu_1), \nu_2\}$, other than the element in the position corresponding to that of ν_1^* in V_1 . For example, if $\nu_1 = 0.15$ and $\nu_2 = 0.4$, then the following pairs of break-points also have the same critical values: $(\nu_1^*, \nu_2^*) = (0.15, 0.75); (0.25, 0.85); (0.25, 0.4); (0.6, 0.85); (0.6, 0.75)$.

The R code is provided in the appendix, and a “commented” version of it and the Eviews program can be downloaded from <http://web.uvic.ca/~dgiles/downloads/johansen/index.html>. A selection of the critical values for the case of a single structural break ($q = 2$), and for $1 \leq (p - r) \leq 10$, appears in Table 1. This upper bound on $(p - r)$ is suggested by Johansen *et al.* (2000, p.230). A more complete table for $q = 2$, and tables of critical values for the case of two structural breaks ($q = 3$) are available in an Excel workbook that can be downloaded from the URL noted above. Running the code with $q = 1$ generates asymptotic p -values and critical values for the no-breaks case.

4. Conclusions

Modified tests for multivariate cointegration when the data exhibit structural breaks have been proposed by Johansen *et al.* (2000). This testing problem arises frequently in empirical economic analysis, and is of considerable importance to practitioners. The asymptotic distributions of these modified cointegration test statistics differ from that of Johansen’s (1988) usual “Trace test”. Although Johansen *et al.* (2000) provide information that allows one to compute appropriate asymptotic critical values, this is no simple matter for the typical practitioner. Also of practical interest is a simple way of calculating asymptotic p -values for these tests.

We supply code that facilitates the easy calculation of both p -values and critical values for these modified cointegration tests, and we provide access to tables of critical values for a selection of situations. Hopefully, this will encourage a more widespread use of these valuable cointegration tests.

References

Byrne, J. P. and R. Perman (2007). Unit roots and structural breaks: A survey of the literature. In B. B. Rao (ed.), *Cointegration for the Applied Economist*, Second Ed.. Palgrave Macmillan, New York, 129-142.

Glynn, J., N. Perera and R. Verma (2007). Unit root tests and structural breaks: A survey with applications. *Revista de Métodos Cuantitativos Para la Economía y la Empresa*, 3, 63–79.

IHS EViews (2011). *EViews 7.2*, Irvine CA.

Inoue, A. (1999). Tests of cointegrating rank with a trend-break. *Journal of Econometrics*, 90, 215–237.

Johansen, S. (1988). Statistical analysis of cointegration vectors. *Journal of Economic Dynamics and Control*, 12, 231–254.

Johansen, S., R. Mosconi and B. Nielsen (2000). Cointegration analysis in the presence of structural breaks in the deterministic trend. *Econometrics Journal*, 3, 216-249.

Joyeux, R. (2007). How to deal with structural breaks in practical cointegration analysis? In B. B. Rao (ed.), *Cointegration for the Applied Economist*, Second Ed.. Palgrave Macmillan, New York, 195-221.

Perron, P. (2006). Dealing with structural breaks. In T. C. Mills and K. Patterson (eds.), *Palgrave Handbook of Econometrics, Vol. 1: Econometric Theory*. Palgrave Macmillan, New York, 278-352.

R Development Core Team (2011). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org/>.

Table 1: Asymptotic critical values

$(p - r)$	$H_l(r)$			$H_c(r)$		
	0.9	0.95	0.99	0.9	0.95	0.99
$\nu_1 = 0.1$						
1	13.7	15.8	20.2	9.9	11.7	15.7
2	28.8	31.6	37.2	22.0	24.5	29.5
3	47.9	51.4	58.2	38.2	41.3	47.5
4	70.9	75.0	82.9	58.2	61.9	69.3
5	98.0	102.6	111.8	82.2	86.5	95.1
6	129.0	134.3	144.6	110.1	115.1	124.9
7	164.2	170.1	181.6	142.0	147.7	158.6
8	203.5	210.0	222.7	178.0	184.2	196.3
9	247.6	254.8	268.5	218.5	225.4	238.6
10	297.6	305.3	320.1	264.6	272.1	286.3
$\nu_1 = 0.2$						
1	15.0	17.2	21.8	10.5	12.3	16.9
2	31.3	34.2	40.0	23.1	25.5	30.7
3	51.5	55.0	62.0	39.5	42.6	48.9
4	75.5	79.6	87.8	59.8	63.5	71.0
5	103.6	108.4	117.7	84.1	88.5	97.1
6	135.7	141.2	151.7	112.3	117.4	127.2
7	172.0	178.0	189.7	144.5	150.2	161.2
8	212.4	219.0	231.9	180.8	187.1	199.3
9	257.5	264.8	278.8	221.6	228.5	241.8
10	308.5	316.4	331.5	268.1	275.6	289.9
$\nu_1 = 0.3$						
1	16.1	18.2	22.9	10.8	12.6	16.6
2	33.1	36.1	42.0	23.6	26.1	31.3
3	53.8	57.4	64.6	40.1	43.2	49.6
4	78.4	82.6	90.9	60.3	64.1	71.6
5	106.9	111.8	121.3	84.5	89.0	97.6
6	139.4	144.9	155.6	112.7	117.7	127.6
7	175.9	182.0	193.9	144.7	150.4	161.5
8	216.5	223.2	236.3	180.8	187.1	199.3
9	261.6	269.0	283.2	221.4	228.3	241.6
10	312.5	320.5	335.8	267.6	275.1	289.4

Table 1 (continued)

$(p - r)$	$H_l(r)$			$H_c(r)$		
	0.9	0.95	0.99	0.9	0.95	0.99
$\nu_1 = 0.4$						
1	16.7	18.8	23.3	11.0	12.8	16.7
2	34.2	37.1	43.1	23.8	26.3	31.5
3	55.1	58.8	66.0	40.2	43.4	49.8
4	79.8	84.1	92.5	60.3	64.1	71.7
5	108.4	113.3	122.8	84.4	88.8	97.5
6	140.8	146.4	157.1	112.3	117.4	127.2
7	177.1	183.3	195.2	144.1	149.8	160.8
8	217.3	224.1	237.2	179.9	186.2	198.4
9	261.9	269.3	283.6	220.1	227.0	240.3
10	312.0	320.0	335.4	265.9	273.4	287.7
$\nu_1 = 0.5$						
1	16.8	18.9	23.3	11.1	12.8	16.7
2	34.5	37.4	43.4	23.9	26.4	31.7
3	55.5	59.1	66.3	40.3	43.5	49.9
4	80.1	84.4	92.8	60.4	64.2	71.7
5	108.6	113.5	123.1	84.4	88.8	97.5
6	140.8	146.4	157.1	112.3	117.3	127.2
7	176.5	182.9	194.9	144.1	149.7	160.8
8	216.5	223.3	236.4	179.8	186.1	198.3
9	260.4	267.9	282.2	220.0	226.9	240.2
10	309.7	317.7	333.1	265.8	273.2	287.5

Note: Critical values for $(1 - \nu_1)$ are the same as those for ν_1 .

Appendix: R Code

```
# Computes asymptotic p-values and critical values for Johansen et al.
#(2000) modified Trace tests for cointegration in presence of structural
# breaks
# Authors: Ryan T. Godwin & David E. Giles, June 2011
# Download From: http://web.uvic.ca/~dqfiles/downloads/johansen/index.html
#=====
library(gplots)
# User has to assign the following 3 values
#=====
q <- 3      # q=1 implies no breaks & the values of v1 & v2 are ignored
v1 <- 0.2
v2 <- 0.85
#=====
# If user requires p-value for one or both Trace statistics, alter one or
# both of the next 2 lines
traceL <- 0      # Value of Hl(r) statistic
traceC <- 0      # Value of Hc(r) statistic
#=====
pr_max <- 10    # DO NOT make (p-r) greater than 10 - see Johansen et al.
a <- c(0, 0, min(v1, v2-v1, 1-v2)) [q]
b <- c(0, min(v1, 1-v1), median(c(v1,v2-v1,1-v2))) [q]
pr <- c(1:pr_max)
lmL <- 3.06+0.456*pr+1.47*a+0.993*b-0.0269*pr^2-0.0363*a*pr-0.0195*b*pr
  -4.21*a^2-2.35*b^2+0.000840*pr^3+6.01*a^3-1.33*a^2*b+2.04*b^3-2.05/pr
  -0.304*a/pr+1.06*b/pr+9.35*a^2/pr+3.82*a*b/pr+2.12*b^2/pr-22.8*a^3/pr
  -7.15*a*b^2/pr-4.95*b^3/pr+0.681/pr^2-0.828*b/pr^2-5.43*a^2/pr^2
  +13.1*a^3/pr^2+1.5*b^3/pr^2
lvL <- 3.97+0.314*pr+1.79*a+0.256*b-0.00898*pr^2-0.0688*a*pr-4.08*a^2
  +4.75*a^3-0.587*b^3-2.47/pr+1.62*a/pr+3.13*b/pr-4.52*a^2/pr-1.21*a*b/pr
  -5.87*b^2/pr+4.89*b^3/pr+0.874/pr^2-0.865*b/pr^2
meanL <- exp(lmL) - (3-q)*pr
varL <- exp(lvL) - 2*(3-q)*pr
thetaL <- varL/meanL
kL <- meanL^2/varL
lmC <- 2.80+0.501*pr+1.43*a+0.399*b-0.0309*pr^2-0.0600*a*pr-5.72*a^2
  -1.12*a*b-1.70*b^2+0.000974*pr^3+0.168*a^2*pr+6.34*a^3+1.89*a*b^2
  +1.85*b^3-2.19/pr-0.438*a/pr+1.79*b/pr +6.03*a^2/pr+3.08*a*b/pr
  -1.97*b^2/pr-8.08*a^3/pr-5.79*a*b^2/pr+0.717/pr^2-1.29*b/pr^2
  -1.52*a^2/pr^2+2.87*b^2/pr^2-2.03*b^3/pr^2
lvC <- 3.78+0.346*pr+0.859*a-0.0106*pr^2-0.0339*a*pr-2.35*a^2+3.95*a^3
  -0.282*b^3-2.73/pr +0.874*a/pr+2.36*b/pr-2.88*a^2/pr-4.44*b^2/pr
  +4.31*b^3/pr+1.02/pr^2-0.807*b/pr^2
```

```

meanC <- exp(lmC) - (3-q)*pr
varC <- exp(lvC) - 2*(3-q)*pr
thetaC <- varC/meanC
kC <- meanC^2/varC
windows(6,3.8)
crit <- cbind(sapply(c(.9,.95,.99) , function(x)
  sprintf("%.2f",round(c(qgamma(x, shape=kL,scale=thetaL),2))),
  sapply(c(.9,.95,.99) , function(x) sprintf("%.2f",round(c(qgamma(x,
  shape=kC,scale=thetaC),2))))))
colnames(crit) <- rep(c(0.9,0.95,0.99),2)
rownames(crit) <- pr
textplot(crit,cex=1)
text(.064,.91,"p-r",font=2)
text(.345,1,expression(paste(plain(H)[1],"(r) test")),col=2)
text(.821,1,expression(paste(plain(H)[c],"(r) test")),col=4)
title("Asymptotic Critical Values")
if(traceL!=0){
windows(4,3.8)
pvalL <- matrix(sprintf("%.3f",round(1 - pgamma(traceL, shape=kL, scale=
  thetaL),3)))
rownames(pvalL) <- pr
textplot(pvalL,cex=1,show.colnames=F)
text(.69,.96,substitute(paste("Pr(",plain(H)[1],">",nn,")")),list(nn=traceL)
),col=2)
text(.45,.96,"p-r",font=2)
title("Asymptotic p-Values")
}
if(traceC!=0){
windows(3,3.8)
pvalL <- matrix(sprintf("%.3f",round(1 - pgamma(traceC, shape=kC, scale=
  thetaC),3)))
rownames(pvalL) <- pr
textplot(pvalL,cex=1,show.colnames=F)
text(.78,.96,substitute(paste("Pr(",plain(H)[c],">",nn,")")),list(nn=traceC)
),col=4)
text(.43,.96,"p-r",font=2)
title("Asymptotic p-Values")
}

```