INFORMATION FOR SUBSCRIBERS

Oxford Bulletin of Economics & Statistics is published four times a year, in February, May, August and November by *Basil Blackwell Ltd*, 108 Cowley Road, Oxford OX4 1JF or PO Box 1320, Murray Hill Station, NY 10156, USA.

SUBSCRIPTION PRICES 1989	UK	USA	OVERSEAS incl. EIRE
Institutions	£43.00	US$91.00	£53.00
Individuals	£22.00	US$47.00	£27.25
Single issues	£11.25	US$22.75	£13.75

Back issues are available from the publisher at the current single copy rate.

US mailing. Second class postage paid at New York. Postmaster; send address corrections to *Oxford Bulletin of Economics & Statistics*, c/o Expediters of the Printed Word Ltd, 515 Madison Avenue, New York 10022, USA — (US mailing agent).

Photocopying. Permission is granted by the publisher for those registered with the Copyright Clearance Center to authorize internal photocopying of any article in the journal for a fee of $3.00 per copy per article. Payment should be made direct to Copyright Clearance Center, 21 Congress Street, Salem, Massachusetts 01970, USA. Individuals or institutions not registered with the CCC, must obtain permission to photocopy either from the CCC or from Basil Blackwell, 108 Cowley Road, Oxford OX4 1JF, when the same fees will apply.

Advertising. For details contact the Advertising Manager or write c/o the publishers.

ISSN 0305-9049 Printed and bound in Great Britain by Page Bros Ltd, Norwich

NOTES TO CONTRIBUTORS

1. Three copies of any submission are required and should be sent to the BULLETIN Editorial Board, Institute of Economics and Statistics, St. Cross Building, Manor Road, Oxford OX1 3UL. Copies should be typed double-spaced on one side of the paper only and record the author's name in the form he wishes it to be printed.

2. Diagrams should be in a form-suitable for immediate reproduction; photocopies are unsatisfactory. They should be accompanied by the basic statistics that were required for the preparation, which will not be published.

3. Contributors should provide supplementary mathematical working and supplementary notes on methods and data, which will not be published.

4. Mathematical notation: Contributors are asked to conform to the conventions adopted in recent articles published in the BULLETIN. Avoid unusual symbols where common ones will do. Avoid fractions in the form $\frac{a}{b}$ when a/b will serve. Use special care in distinguishing symbols you cannot type, and avoid letter symbols likely to cause confusion (e.g. 1 and 0). A list of symbols should be provided in respect of articles where extensive use is made of different symbols. This will not be published. Avoid circumflexes, bars, tildes and dots above letters where possible. Check the layout of mathematical expressions especially alignment, centring, length of fraction lines, and type, position and closure of brackets.

(*continued on inside back cover*)

OXFORD BULLETIN OF ECONOMICS AND STATISTICS, 51, 2 (1989)
0305-9049 $3.00

PRACTITIONERS CORNER

Maximum Likelihood Estimation of Cointegration Vectors: An Example of The Johansen Procedure‡

S. G. Hall

INTRODUCTION

The concept of cointegration has attracted increasing attention over recent years, the key paper of Engle and Granger (1987) providing a spur to a great deal of research in a number of directions. One strand of this research involves the estimation and testing of cointegrating vectors within an OLS framework. A two-step estimator may be constructed, by first estimating a static regression, then taking the residuals from this equation and including them in a full dynamic model. Engle and Granger show that this estimator is consistent and that the convergence properties of the parameter estimates are superior to standard OLS. This procedure involves normalizing the cointegrating vector on one of the variables which makes the assumption that the corresponding element of the cointegrating vector is non-zero. Doubts have been raised about the usefulness of this procedure in the light of a considerable degree of small sample bias which may occur in the parameter estimates (e.g. Banerjee *et al.* (1986)). Despite this difficulty, a number of applications of the two step estimator have been performed, including Hall (1986), Jenkinson (1986) and Drobney and Hall (1987).

In practical applications there are a number of disadvantages to the two-step procedure which are perhaps more important than the small sample bias. In particular two problems are largely unresolved; first the assumption is made that the cointegrating vector is unique, this may not, however, be the case and the two-step procedure provides no framework for addressing this question. Second, the test procedures do not have well defined limiting distributions and as a result testing for cointegration is not a straightforward procedure.

A recent paper by Johansen (1988) suggests a maximum likelihood estimation procedure which offers solutions for both of these problems. It provides

‡I would like to thank D. F. Hendry and S. Johansen for helpful comments on an earlier draft of this paper, any remaining errors are of course my own responsibility. The views expressed in this paper do not necessarily represent those of the Bank of England.

estimates of all the cointegrating vectors which exist between a set of variables as well as test statistics for the number of cointegrating vectors which have an exact limiting distribution which is a function of only one parameter. In this paper the model of Hall (1987) will be estimated for aggregate UK data using this new procedure. The maximum likelihood estimates of the cointegrating vector may then be compared with those derived by the Engle and Granger two-step procedure.

The next section of this paper will give an account of the Johansen estimation techniques, the third section will provide an application of this technique to UK wage data, the fourth section will draw some comparisons and conclusions.

THE JOHANSEN PROCEDURE

Johansen sets his analysis within the following framework. Begin by defining a general polynomial distributed lag model of a vector of variables X as

$$X_t = \pi_1 X_{t-1} + \ldots + \pi_k X_{t-k} + \varepsilon_t \quad t = 1, \ldots, T \tag{1}$$

where X is a vector of N variables of interest; and ε_t is an independently identically distributed N dimensional vector with zero mean and variance matrix Ω. Within this framework the long run, or cointegrating matrix is given by

$$I - \pi_1 - \pi_2 \ldots - \pi_k = \pi \tag{2}$$

π will therefore be an NXN matrix and the number of distinct cointegrating vectors which exist between the variables of X, r, will be given by the rank of π. In general if X consists of variables which must be differenced once in order to be stationary (integrated of order one or $I(1)$) then, at most, r must be equal to $N-1$, so that $r \leq N-1$. Now we define two matrices α, β both of which are Nxr such that

$$\pi = \alpha\beta'$$

and so the rows of β form the r distinct cointegrating vectors.

Johansen then demonstrates the following Theorem.

Theorem: The maximum likelihood estimate of the space spanned by β is the space spanned by the r canonical variates corresponding to the r largest squared canonical correlations between the residuals of X_{t-k} and ΔX_t corrected for the effect of the lagged differences of the X process. The likelihood ratio test statistic for the hypothesis that there are at most r cointegrating vectors is

$$-2 \ln Q = -T \sum_{i=r+1}^{N} \ln(1 - \hat{\lambda}_i) \tag{3}$$

where $\hat{\lambda}_{r+1} \ldots \hat{\lambda}_N$ are the $N-r$ smallest squared canonical correlations. Johansen then goes on to demonstrate the properties of the maximum likelihood estimates and, more importantly, he shows that the likelihood ratio test has an asymptotic distribution which is a function of an $N-r$ dimensional Brownian motion which is independent of any nuisance parameters. This means that a set of critical values can be tabulated which will be correct for all models. He demonstrates that the space spanned by β is consistently estimated by the space spanned by $\hat{\beta}$.

In order to implement this Theorem we begin by reparameterizing (1) into the following error correction model.

$$\Delta X_t = \Gamma_1 \Delta X_{t-1} + \quad + \Gamma_{k-1} \Delta X_{t-k+1} + \Gamma_k X_{t-k} + \varepsilon_t \tag{4}$$

where

$$\Gamma_i = -I + \pi_1 + \ldots \pi_i;\ i = 1 \ldots k$$

The equilibrium matrix π is now clearly identified as $-\Gamma_k$.

Johansen's suggested procedure begins by regressing ΔX_t on the lagged differences of ΔX_t and defining a set of residuals R_{0t}, then regressing X_{t-k} on the lagged differences and defining R_{kt}. The likelihood function, in terms of α, β and Ω is then proportional to

$$L(\alpha, \beta, \Omega) = |\Omega|^{-T/2} \exp\left[-\frac{1}{2} \sum_{t=1}^{T} (R_{ot} + \alpha\beta' R_{kt})' \Omega^{-1} (R_{ot} + \alpha\beta' R_{kt}) \right] \tag{5}$$

If β were fixed we could maximize over α and Ω by a regression of R_{0t} on $-\beta' R_{kt}$ which gives

$$\hat{\alpha}(\beta) = -S_{0k}\beta(\beta' S_{kk}\beta)^{-1} \tag{6}$$

and

$$\hat{\Omega}(\beta) = S_{00} - S_{0k}\beta(\beta' S_{kk}\beta)^{-1}\beta' S_{ko} \tag{7}$$

where

$$S_{ij} = T^{-1} \sum_{t=1}^{T} R_{it} R'_{jt} \quad i, j = 0, k$$

and so maximizing the likelihood function may be reduced to minimizing

$$|S_{00} - S_{0k}\beta(\beta' S_{kk}\beta)^{-1}\beta' S_{k0}| \tag{8}$$

and it may be shown that (8) will be minimized when

$$|\beta' S_{kk}\beta - \beta' S_{k0} S_{00}^{-1} S_{0k}\beta| / |\beta' S_{kk}\beta| \tag{9}$$

attains a minimum with respect to β.

We now define a diagonal matrix D which consists of the ordered eigenvalues $\lambda_1 > \ldots > \lambda_N$ of $S_{k0} S_{00}^{-1} S_{0k}$ with respect to S_{kk}. That is λ_i satisfies

$$|\lambda S_{kk} - S_{k0} S_{00}^{-1} S_{0k}| = 0 \tag{10}$$

Define E to be the corresponding matrix of eigenvectors so that

$$S_{kk}ED = S_{k0}S_{00}^{-1}S_{0k}E \qquad (11)$$

where we normalize E such that $E'S_{kk}E = I$.

The maximum likelihood estimator of β is now given by the first r rows of E, that is, the first r eigenvectors of $S_{k0}S_{00}^{-1}S_{0k}$ with respect to S_{kk}. These are the canonical variates and the corresponding eigenvalues are the squared canonical correlations of R_k with respect to R_0. These eigenvalues may then be used in the test proposed in (3) to test either for the existence of a cointegrating vector $r = 1$ or the number of cointegrating vectors $N > r > 1$.

AN APPLICATION TO UK WAGE RATES

This section will extend the work reported in Hall (1986) using this new technique. That application considered the cointegrating vectors which existed between the log of wages, LW, the log of prices LP, the log of productivity $LPRODS$, the log of average hours worked $LAVH$ and the percentage rate of unemployment UPC. It was shown that both LW and LP were $I(2)$ series which cointegrated to produce an $I(1)$ series for the real wage LRW. In this work we will confine the model to consider $I(1)$ variables only by working with the real wage rather than nominal wages and prices. This gives a vector of four variables LRW, $LPRODS$, $LAVH$ and UPC.

Given that the data is quarterly and that the lag structure reported in Hall (1986) is fairly simple, a polynomial distributed lag model was specified with a maximum lag of four quarters. So k in (1) takes the value 4.

The procedure then begins by performing two sets of 4 regressions, in the first set we regress the first difference in each variable on all the lagged differences of all the variables up to the third lag. So the first regression will be

$$\Delta LRW = a_0 + \sum_{i=1}^{3} (b_i \Delta \mathrm{LPRODS}_{t-i} + c_i \Delta LAVH_{t-i} + d_i \Delta UPC_{t-i} + e_i \Delta LRW_{t-i}) \qquad (12)$$

The residuals from each of these regressions are stored and they form the R_0 residuals.

The next step involves another four regressions of the 4 period lagged level of each variable on the same set of lagged first difference terms as given in (10). The residuals from these regressions then form the set of R_k residuals. These two sets of residuals may then be used to form the four matrices S_{00}, S_{0k}, S_{kk} and S_{k0} defined in (7). Finally the eigenvalues and eigenvectors of (10) may be calculated.

On a point of computation, (10) is a non-standard form of the eigenvalue problem which is sometimes referred to as the generalized eigenvalue problem. It may be solved directly by some advanced programming libraries

such as the Nag library or it may be put into a more tractable form using the following suggestions.[1]

Decompose the matrix $S_{kk} = CC'$ using a Cholesky decomposition. Then note that the eigenvalues for (10) will also solve

$$|\lambda I - C^{-1}S_{k0}S_{00}^{-1}S_{0k}C'^{-1}| = 0$$

This is now a standard eigenvalue problem involving a symmetric positive definite matrix $C^{-1}S_{k0}S_{00}^{-1}S_{0k}C'^{-1}$, which shows that the eigenvalues of (10) must be real and positive. Now if $V_1 \dots V_N$ denote the normalized eigenvectors such that $V'V = I$ then $E = C'^{-1}V$ will give the eigenvectors of the original problem normalized such that $E'S_{kk}E = I$.

The results of this procedure are reported below in Table 1 (estimation was performed using the authors regression program $REG-X$).

The test statistics that there are at most r cointegrating vectors are given by

r	LR test	95% Critical value[2]
0	66.8	38.6
1	23.7	23.8
2	11.6	12.0
3	1.41	4.2

The likelihood ratio test that there are at most 0 cointegrating vectors easily rejects the null hypothesis, so we know that there is at least one cointegrating vector. The likelihood ratio tests that there are at most 1 or 2 cointegrating vectors are very close to the 95 percent critical values and the test that there are at most three cointegrating vectors is well within the rejection criteria. So there is clearly one cointegrating vector, there is some doubt, based on the likelihood ratio tests, that there may be more than one cointegrating vector although the coefficients of the second and third cointegrating vectors do not seem to make economic sense.

If we renormalize the values of the first eigenvector so that real wages takes the value -1 we get

Real wages	Productivity	Unemployment	Hours
−1	1.099	−0.562	−0.94

It is interesting to compare this eigenvector with table 3 of Hall (1986). That table estimated the coefficients of the cointegrating vector in all its inversions. The parameter on $LPRODS$ varied between 0.85 and 1.21, its ML estimate in Table 1 is 1.099. The parameter on UPC varied between -0.56 and -3.52 with four of the five estimates being greater than -0.73; the ML estimator is -0.562. Finally the parameter of $LAVH$ varies between -1.65 and -2.64 and the ML estimator is -0.94. In this paper we are working with total wage rates rather than hourly wage rates so the corresponding

[1] I am grateful to Professor Johansen for suggesting this procedure.
[2] Critical values are taken from the tables in Johansen (1988).

hourly effect would be $-0.65\ldots-1.64$. In every case the *ML* estimator therefore lies within the space defined by the different OLS estimates of the cointegrating vector.

If the cointegrating vector is not unique then the second vector would be estimated by the eigenvector corresponding to the eigenvalue 0.19. The vector would seem to offer no convincing economic interpretation, although given the small size of the real wage effect it may be suggesting a relationship between unemployment and productivity.

CONCLUSION

The *ML* estimator has been shown to provide estimates of the cointegrating vector which conform well with those given by OLS. It is also clear that different versions of the OLS equations are not providing estimates of different cointegrating vectors and that the differences are due to the small sample bias of the OLS estimates which should disappear asymptotically.

Bank of England

Date of Receipt of Final Manuscript: August 1988

REFERENCES

Banerjee, A., Dolado, J. J., Hendry, D. F. and Smith, G. W. (1986). 'Explaining Equilibrium Relationships in Econometrics through Static Models: Some Monte Carlo Evidence', *BULLETIN*, Vol. 48, No. 3, pp. 253–78.

Drobny, A. and Hall, S. G. (1987). 'An Investigation of the Long Run Properties of Aggregate Non-durable Consumers Expenditure in the UK', NIESR Discussion Paper No. 128.

Engle, R. F. and Granger, C. W. J. (1987). 'Co-integration and Error Correction: Representation, Estimation and Testing', *Econometrica*, Vol. 55, pp. 251–76.

Hall, S. G. (1986). 'An Application of the Granger and Engle Two-step Estimation Procedure to United Kingdom Aggregate Wage Data', *BULLETIN*, Vol. 48, No. 3, pp. 229–40.

Jenkinson, T. J. (1986). 'Testing Neo-Classical Theories of Labour Demand: An Application of Cointegration Techniques', *BULLETIN* Vol. 48, No. 3, pp. 241–51.

Johansen, S. (1988). 'Statistical Analysis of Cointegration Vectors', *Journal of Economic Dynamics and Control*.

(*continued from inside front cover*)

5. Statistical tables should be clearly headed and be as self-explanatory as possible. Units of measurement, base-dates for index numbers, geographical area covered and sources should be clearly stated. Whenever they feel that the referee would have difficulty in testing the derivation of their statistics, contributors should provide supplementary notes on the methods used, which will not be published.

6. If a bibliographical list is given it should conform to the convention adopted in recent articles published in the BULLETIN.

7. Contributors are responsible for the correction of proofs. Printer's errors should be marked in red and author's changes in blue. Alterations must be kept down to a very low level. The editors reserve the right to veto and/or charge for alterations made in proof. Proofs should be returned within two weeks (three weeks for overseas contributors).

8. Articles should contain a final section in which the author summarizes and concludes in a way which will be intelligible to the non-specialist reader. A separate abstract for the *Journal of Economic Literature*, not exceeding 100 words, should accompany a submission.

9. Contributors receive 25 free off-prints. Further off-prints may be ordered at the contributor's own expense.

10. Anyone wanting to submit comments on a BULLETIN article should start by sending a copy to the author, inviting full discussion. The commentator is asked to allow the author a reasonable time to reply before he sends anything to the editors, and to enclose any reply which he may have received from the author. If the editors accept the comment for publication, they will send a copy to the original author and give him a chance to make a short reply in the same issue of the BULLETIN.