JOURNAL OF APPLIED ECONOMETRICS, VOL. 10, 205–210 (1995)

SOFTWARE REVIEWS

REVIEW OF Coint 2.0

SERENA NG
CRDE and Department of Economics, University of Montreal, CP 6128 Succ. Centre Ville, Montreal, Quebec H3C 3J7, Canada

1. WHAT IS IN Coint 2.0?

Coint 2.0 by Predicta Software is a comprehensive set of econometrics procedures written in Gauss for testing and estimating linear dynamic models with non-stationary variables. The package consists of seven modules: UNIT.SRC, CREGRS.SRC, BASE.SRC, LRVAR.SRC, BAYES.SRC, ARMA.SRC, and KERNELS.SRC. I find it more convenient to think of the contents of Coint 2.0 in terms of their tasks.

Hypothesis Testing

(1) Unit root tests
- (a) The Phillips (1987) and Phillips and Perron (1988) Z_α and Z_t tests
- (b) The Augmented Dickey–Fuller (ADF) test of Said and Dickey (1984)
- (c) The G and J statistics of Park and Choi (1988)
- (d) The posterior distribution for a unit root in an AR(p) model with deterministic trend using a Jeffreys or a uniform prior on the regression coefficients as suggested in Phillips (1991a,b).

(2) Cointegration tests
- (a) The residual-based Z_α and Z_t tests
- (b) The residual-based ADF test
- (c) The Stock and Watson (1988) common trend statistic
- (d) The PU and PZ statistics of Phillips and Ouliaris (1990)
- (e) The H(p, q) statistic of Park (1990)
- (f) The posterior distribution for a unit root in the residuals of a cointegrating regression using a Jeffreys prior, an e-prior, or a uniform prior as discussed in Zivot and Phillips (1994)
- (g) The MeanF and SupF tests for the stability of cointegrating vectors of Hansen (1991).

CCC 0883–7252/95/020205–06

Received January 1995
Revised January 1995

Estimation of Cointegrating Regressions

(1) Time domain
  (a) The CCR (Canonical Cointegrating Regressions) of Park (1992)
  (b) The FM (Fully Modified) estimator of Phillips and Hansen (1990)
  (c) The FM-OLS and FM-VAR of Phillips (1993)
  (d) The FM-GIVE and FM-GMM of Kitamura and Phillips (1992)
  (e) The ML estimator of Johansen (1988)
  (f) The VECM estimator of Saikkonen (1991)
(2) Frequency domain
  (a) The full bandwidth and frequency zero spectral estimators of Phillips (1990)
  (b) The full bandwidth and frequency zero GIVE spectral estimators of Phillips (1990)

Supplementary Procedures

Model selection
The procedures select the optimal order of p, q, and tr in an ARMA(p,q) with polynomial trend of order tr using the AIC (Akaike), the BIC (Schwarz), and the PIC (Posterior) criteria discussed in Phillips and Ploberger (1994).

Estimation of stationary ARMA models
The procedures estimate ARMA(p,q,tr) models using the two- or the three-stage Recursive Least Squares Estimator of Hannan and Rissanen (1982) and Kavelieris (1991) for given p, q, tr, and r, where r is the order of the autoregression in the first stage. Additional procedures allow the user to use the AIC to choose r, or use the BIC to choose the optimal values of p, q, and tr from some pre-specified *pmax*, *qmax*, *tr-max* and r.

Estimation of the spectral density function
The procedures perform parametric and non-parametric estimations of the spectral density function at the zero and other user-specified frequencies. The parametric estimates are based on fitting an ARMA(p,q) model, where p and q can be predetermined or chosen by information criteria. The non-parametric estimates are kernel based and Coint 2.0 offers a choice of 11 windows. The user can always preset the truncation lag, but for five of the kernels, the data-dependent method of Andrews (1991) can be used to select the optimal lag length. The user can also opt to prewhiten the data using the method of either Andrews and Monahan (1991) or Lee and Phillips (1993).

2. USING Coint 2.0

Coint 2.0 works with Gauss 3.0 and later releases. The resource limits are completely determined by Gauss. For frequently used economic time series data, 4 MB of workspace memory should be more than adequate, but more memory may be required for efficient use of procedures which optionally invoke Gauss's graphical routines. Coint 2.0 was developed under the DOS environment. I tested Coint 2.0 on a 486DX/33 mhz PC running a DOS box in OS/2 2.1 and encountered no operating-system related problem. I was informed that a UNIX version of Coint 2.0 will soon be available.

Installing Coint 2.0 requires nothing more than copying the .src files and one .lcg file into the appropriate directories and loading up the COINT and PGRAPH libraries at startup. Example files are also provided but need not be copied to the hard disk. The package comes with a

80-page bound manual which explains the syntax for using the procedures. The manual is generally well written. Several similarly named procedures (e.g. ARMABC and ARMABIC3) are functionally identical except that one has an option to graph the results. The manual is laid out as though these are separate procedures, which is somewhat confusing. I have also found some mistakes in the manual. Some are purely typographical; others are the result of rewriting the procedures from 1.0 without changing the manual. I was assured that these problems will be fixed in the next release.

The mathematical formulae for the statistics and estimators are not described in the manual but complete references are provided. The instructions are written concisely and are in most cases intuitively obvious. Further elaboration of the econometrics in the manual will most likely turn it into a textbook, which is not the purpose of the package. I do find, however, that when a procedure is based on an unpublished or a lesser-known article, I would have liked to have the definition of the statistic appear in the manual just to be sure that the input parameters are entered correctly.[1] This said, a user who has read the cited papers should have no difficulty executing the procedures and interpreting the results.

A typical hypothesis testing procedure would require the user to input (1) the order of the deterministic trend function; (2) the order of the autoregression, and (3) the number of lags used to compute the long-run variance. The choice of the kernel, whether or not to do prewhitening, and whether or not to use the automatic selection of the bandwidth are controlled through the global variables, _ker_fun, _filter, and _aband respectively. For example, to test if y and x are stochastically cointegrated around a linear trend using the Z_α and Z_t statistics with four lags of autocovariances weighted by the Parzen window and without prewhitening the residuals, the commands:

```
p = 4;                 /* number of lagged autocovariances */
d = 1;                 /* include a first order polynomial trend in the cointegrating regression */
ker_fun = &parzen;     /* use the Parzen kernel for smoothing the autocovariances */
_aband = 0;            /* disable automatic bandwidth selection */
_filter = 0;           /* disable prewhitening of residuals */
{alpha, xza, xzt, cv_za, cv_zt} = cza(y,x,p,d)
```

would return the autoregressive coefficient (alpha), the statistics (xza and xzt), and the critical values (cv_za and cv_zt). The user can then print the output in the format he/she sees fit. Gauss would flag an error if the name of an input or output variable conflicts with the name of a local variable or procedure, but the solution to this is obvious. Coint 2.0 provides critical values at the upper and lower 1%, 5%, and 10% levels for models with polynomial trends of up to fifth order and with up to five regressors. However, for commonly used models with linear or quadratic trends, more accurate critical values can be found in MacKinnon (1991, 1994).

The syntax for using the estimation procedures is similar to that of the hypothesis-testing procedures with one exception. The estimation procedures accommodate deterministic trends that may not be polynomial time trends. The user is required to supply a matrix of deterministic terms to the estimation procedures whether or not deterministic components are to be included in the regression. The procedures then check the status of the global variable _nodet to see if use of these deterministic terms is to be suppressed. Hence to estimate the cointegrating vector

[1] For example, the user needs to input a 'specified AR lag' and an 'AR lag for long AR in the first stage' to do a recursive three-stage estimation of an ARMA model. The meaning of these terms are immediate once we understand the context in which they are used.

between y and x with no deterministic component using the Phillips (1993) Fully Modified OLS estimator with long-run variance constructed using the Quadratic kernel with prewhitening and automatic bandwidth selection, the commands:

```
ker_func = &qs;        /* use the quadratic kernel */
_filter = 1;           /* enable prewhitening */
_aband = 1;            /* enable automatic bandwidth selection */
_nodet = 1;            /* suppress deterministic components in the regression */
d = ones(rows(y));     /* a dummy vector of ones */
p = 4;                 /* a dummy parameter since automatic bandwidth selection is turned on */
{beta, vcv} = fm_ols(y,x,d,p)
```

would return the fully modified OLS estimates (beta) and the associated variance–covariance matrix (vcv). But note that the values for d and p are essentially ignored because the flag for _nodet is set to bypass the deterministic variables, and the global variable _aband is turned on to perform automatic selection of the bandwidth. Care must be taken to keep track of these global variables, especially when running spectral regressions, since the moment matrix of regressors will be singular if a constant was included as a regressor. It would seem more natural to require the user to pass the deterministic variables to the procedures only when the variables are to be used in the regressions. But since current versions of Gauss do not allow passing a variable number of arguments to a procedure, this inconvenience seems to be the cost one has to pay for the flexibility of allowing deterministic components that are not polynomial in time (such as a breaking trend).

For procedures that perform multiple regressions, the user should carefully keep track of how the estimates are stored. For example, a three-variable FM-VAR with lagged levels of the variables augmented by four lagged first differences of each variable and a constant as regressors would return a beta matrix that is 16×3 and a VCV that is 48×48. In such a case, one is presumably more interested in the estimates on the variables in level (the cointegrating vectors) rather than those corresponding to the transient dynamics. Thus, to do inference, the user would extract components of the VCV that correspond to beta[13 : 15, 1 : 3] since for FM-VAR, the rows of the beta matrix store the estimates of the first differenced variables before the variables in level. However, other procedures such as the PS store the beta matrix in a different order. The user should take note of these minor inconsistencies across procedures.

The user can always predetermine the values for p, q, and tr on occasions when input of these parameters are required but can also use the model-selection procedures to serve as a guide. The model selection procedures and those for estimating the long run variance do not seem to be fully integrated with the estimation and hypothesis-testing procedures. For example, the parametric spectral density estimators and the Lee and Phillips (1993) prewhitening procedure can be used in a standalone estimation of the long-run variance but are not options of the estimation procedures. In a way, this incomplete integration of the various modules is also a blessing since the procedures for model selection and estimation of long-run variance and ARMA models are useful tools for analysing stationary time series data in their own right. A seamless integration would probably make it more difficult to access these supplementary procedures even when appropriate.

3. OVERVIEW

Coint 2. 0 is distributed by:

Predicta Software Inc.
133 Concord Drive
Madison
CT 06433, USA
Fax/phone: +(203) 421-3784

Aptech Systems Inc.
23804 S.E. Kent-Kangley Road
Maple Valley
WA 98038, USA
Phone: +(206) 432-7855
Fax: +(206) 432-7832

The software sells for US $200 per single-user copy but site licenses are also available. Since the procedures are sold as a package, those interested in just one or two procedures may find the package overpriced. The software comes on a 3.5-inch diskette with source code available. Technical support for registered users is available by fax, phone, or email. The authors are extremely helpful in responding to questions. A solution is often available within the day after emailing them the offending code. I have also identified some more involved bugs, and those problems were also fixed promptly.

Coint 2.0 is a highly specialized set of procedures for analysis of non-stationary time series. It is the most complete collection of such procedures that I have seen. Other products exist, notably the CCR package of Ogaki (also written in Gauss) and the Jesuslius CATS package (written in RATS), but the scope of Coint 2.0 is more general than these alternatives. Of the relevant estimators that have appeared in published articles, the DOLS (developed independently by Stock and Watson, 1993, Phillips, 1991c, and Saikkonen, 1991) and the non-linear ECM of Phillips and Loretan (1991) are the only two procedures omitted from Coint 2.0 that come to my mind.

Since Coint 2.0 is an add-on to Gauss, it would not be appropriate to judge it in terms of its interface as one would with RATS or TSP. In any event, the source code is available should the user wish to include or eliminate some output. I think the real strength of Coint 2.0 is in its implementation of the FM (Fully Modified) estimators. These procedures apply corrections to many commonly used estimators so that they will have classical properties even in the presence of integrated regressors. The computations of these correction factors are sometimes quite involved.

There are still rough edges in Coint 2.0, in large part because the theory is still new and practical problems associated with the econometrics are only beginning to surface. But the authors seem ready to provide full support to the product. Those who want to use the frontier tools in their applied work will not be disappointed with Coint 2.0. It puts frontier econometric theory into practice. I certainly find Coint 2.0 to be a welcome addition to my collection of Gauss procedures.

REFERENCES

Andrews, D. W. K. (1991), 'Heteroskedasticity and autocorrelation consistent covariance matrix estimation', *Econometrica*, **59**, 817–858.

Andrews, D. W. K. and C. Monahan (1991), 'An improved heteroskedasticity and autocorrelation consistent covariance matrix estimator', *Econometrica*, **60**, 953–966.

Hannan, E. J. and J. Rissanen (1982), 'Recursive estimation of mixed autoregressive-moving average order', *Biometrika*, **69**, 81–94. Hansen, B. E. (1991), 'Tests of parameter instability in regressions with I(1) processes'. mimeo, University of Rochester.

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

Juselius, K. (1991), 'CATS in RATS: A manual to Cointegration Analysis', Institute of Economics, University of Copenhagen.

Kitamura, Y. and P. C. B. Phillips (1992), 'Fully Modified IV, GIVE and GMM estimation with possibly non-stationary regressors and instruments', Cowles Foundation Discussion Paper 1043, Yale University.

Kavalieris, L. (1991), 'A note on estimating autoregressive-moving average order', *Biometrika*, **78**, 920–922.

Lee, C. C. and P. C. B. Phillips (1993), 'An ARMA-prefiltered estimator for the long-run variance with an application to unit root tests', mimeo, Cowles Foundation, Yale University.

MacKinnon, J. (1991), 'Critical values for cointegration tests', in C. Granger and G. Mizon (eds), *Advanced Texts in Economics*, Oxford University Press, Oxford.

MacKinnon, J. (1994), 'Approximate asymptotic distribution functions for unit-root and cointegration tests', *Journal of Business & Economic Statistics*, **12**, 167–176.

Ogaki, M. (1993), 'CCR: A user guide', Rochester Centre for Economic Research Working Paper 348, University of Rochester.

Park, J. Y. (1990), 'Testing for unit roots and cointegration by variable addition', in T. B. Fomby and G. F. Rhodes (eds), *Advances in Econometrics: Cointegration, Spurious Regressions and Unit Roots*, pp. 107–133, JAI Press, Greenwich, CT.

Park, J. Y. (1992), 'Canonical cointegrating regressions', *Econometrica*, **60**, 119–144.

Park, J. Y. and B. Choi (1988), 'A new approach to testing for a unit root', Working Paper 88-23, Cornell University.

Phillips, P. C. B. (1990), 'Time series regression with a unit root', *Econometrica*, **55**, 277–301.

Phillips, P. C. B. (1990), 'Spectral regression for cointegrated time series', in W. Barnett (ed.), *Nonparametric and Semiparametric Methods in Economics and Statistics*, Cambridge University Press, Cambridge.

Phillips, P. C. B. (1991a), 'To criticize the critics: an objective bayesian analysis of stochastic trends', *Journal of Applied Econometrics*, **4**, 333–364.

Phillips, P. C. B. (1991b), 'Bayesian routes and unit roots: de rebus prioribus semper est disputandum', *Journal of Applied Econometrics*, **6**, 435–474.

Phillips, P. C. B. (1991c), 'Optimal inference in cointegrated systems', *Econometrica*, **59**, 283–306.

Phillips, P. C. B. (1993), 'Fully modified least squares and vector autoregression', Cowles Foundation Discussion Paper No. 1047, Yale University.

Phillips, P. C. B. and B. E. Hansen (1990), 'Statistical inference in instrumental variables regression with I(1) process', *Review of Economic Studies*, **57**, 99–125.

Phillips, P. C. B. and M. Loretan (1991), 'Estimating long-run economic equilibria', Review of Economic Studies, **58**, 407–436.

Phillips, P. C. B. and S. Ouliaris (1990), 'Asymptotic properties of residual based tests for cointegration', *Econometrica*, **58**, 165–193.

Phillips, P. C. B. and P. Perron (1988), 'Testing for a unit root in time series regression', *Biometrika*, **75**, 335–346.

Phillips, P. C. B. and W. Ploberger (1994), 'Posterior odds testing for a unit root with data-based model selection', *Econometric Theory*, **10**, 774–808.

Said, S. E. and D. A. Dickey (1984), 'Testing for unit roots in autoregressive moving average model of unknown order', *Biometrika*, **71**, 599–607.

Saikkonen, P. (1991), 'Asymptotically efficient estimation of cointegrating regressions', *Econometric Theory*, **7**, 1–21.

Schwarz, G. (1978), 'Estimating the dimension of a model', *Annals of Statistics*, **6**, 461–464.

Stock, J. and M. Watson (1988), 'Testing for common trends', *Journal of the American Statistical Association*, **83**, 1097–1107.

Stock, J. and M. Watson (1993), 'A simple estimator of cointegrating vectors in higher order integrated systems', *Econometrica*, **6**, 783–821.

Zivot, E. and P. C. B. Phillips (1994), 'A bayesian analysis of trend determination in economic time series', *Econometric Reviews*, **13**, 291–336.