The RATSletter

Volume 15, No. 1 **July, 2002**

Estima

RATS Version 5.02

RATS Version 5.02 is now available for the PC, Macintosh, and UNIX platforms. This is a minor update which includes a number of bug fixes and several useful new features.

If you have Version 5.0 or 5.01 for Windows or DOS, you can download the update at no charge from our website. You simply need to download and run a patch file that updates your existing copy of RATS. Mac and UNIX can receive a copy of the update via e-mail. To request your update, just e-mail us at *estima@estima.com*. Please be sure to include your name and your RATS serial number in the e-mail.

If you prefer, you can also order a copy of 5.02 on CD, at a cost of $15 (for users who already have Version 5.0 or 5.01). In addition to the updated software, the CD will include updated versions of the manuals in Adobe PDF format, which document the new features described here.

If you are using an older version of RATS, please visit our website or contact us for details on updating.

New BOXJENK Option

The option `REGRESSORS` has been added, which allows extra (non-ARIMA) variables to be added using regression format rather than transfer/intervention style inputs. For instance, the following estimates a REGARIMA model with trading day counts (generated with the new `%TRADEDAY` function) as exogenous variables:

```
dec vect[series] days(7)
do i=1,7
   set days(i) = %tradeday(t,i)
end do i
boxjenk(regress,sdiff=1,ma=1,apply)
logsales
# days
```

The `APPLYDIFFERENCES` option works as it does with other `INPUT` variables.

New DLM Option

The option `Z=VECTOR` or `FRML[VECT]` has been added. This allows for exogenous shifts in the state equation. The state equation specification with **Z**'s permitted is

$$\mathbf{X}_t = \mathbf{A}_t\mathbf{X}_{t-1} + \mathbf{Z}_t + \mathbf{w}_t$$

For instance, if the states are assumed to shrink back towards a fixed vector, the model

$$\mathbf{X}_t = \lambda\mathbf{X}_{t-1} + (1-\lambda)\mu + \mathbf{w}_t$$

can be analyzed with

(continued on page 3)

Other Product News and Reminders

A few other bits of news regarding our RATS software and database products:

We have increased the data-handling capacity of Classroom RATS from 3,000 observations to 6,000 observations, based on feedback from users. Classroom RATS is available for the Mac and Windows platforms. Ordered in quantities of 5 or more, the price is $60 per copy with printed documentation, or $40 copy for a CD only (documentation included in Adobe PDF format).

A reminder that the X11 seasonal adjustment routine is now available as an add-on module for RATS, at a cost of $150. Contact Estima if you would like to add this powerful tool to your copy of RATS 5.

The OECD MEI database continues to expand—the most recent additions include data for India and South Africa. Hungary, Poland, Russia (and former Soviet Republics) , and Mexico were also added in the past couple of years.

Contact Estima for more information on these products.

Panel Data in RATS 5.0

With version 5, we made major changes to the internal handling of panel data sets. The most obvious change was the addition of the instruction **`PREGRESS`**, which automatically handles random and fixed effects regressions, where earlier versions of RATS required the calculations to be done by a sequence of instructions. While implementing this, we also eliminated the restriction that the panel set be *balanced*, that is, RATS 5 no longer requires (for analysis using the packaged panel data procedures) that all individuals have data for exactly the same set of time periods.

Unbalanced panels are not terribly difficult to handle when the effects being allowed are either individual only or time only. However, they are quite tricky when both types of effects are allowed simultaneously. An examination of other statistical packages showed that some simply avoided the problem by not allowing combined effects. When we closely looked at others, we found that the results were often questionable.

(continued on page 4)

Contents

Benchmarking with DLM in RATS

In this context, benchmarking refers to the process of combining high frequency observations on an economic series with lower frequency values for the same series which are considered more reliable. This is a special case of distribution, which is the process of distributing a low frequency value across a higher frequency while preserving the sum.

The IMF has an extensive description of various benchmarking methods in *Quarterly National Accounts Manual—Concepts, Data Sources, and Compilation*, by Bloem, Dippelsman, and Maehle, available on the web at:

http://www.imf.org/external/pubs/ft/qna/2000/textbook/

IMF's preferred method (for most series) is the proportional Denton technique. To summarize, you have a series (*A*) which is observable annually, and a related series (*I*) which is observable quarterly. The annual series is assumed to be in some sense a higher quality series than the higher frequency one. The desire is to distribute the annual number into quarterly values (*X*) which sum to the annual number and which take a quarter to quarter pattern which has some similarity to that taken by the observable quarterly series. The proportional Denton method is to choose a series *X* which minimizes

$$\sum_{t=2}^{T}\left(\frac{X_t}{I_t}-\frac{X_{t-1}}{I_{t-1}}\right)^2$$

subject to the adding up constraint on the *X*'s (that the annual sum of the *X*'s is equal to the observed annual sums).

The simplest way to implement this in RATS is to use Kalman smoothing with **DLM**. The underlying model is quite simple. Define $Z_t = X_t / I_t$. If $Z_t = Z_{t-1} + u_t$, then the objective function is to minimize the sum of u_t^2 subject to the adding up constraint. Creating a state vector which just gives Z_t and three of its lags allows the constraint to be written as a measurement equation with no error:

$$A_t = I_t Z_t + I_{t-1} Z_{t-1} + I_{t-2} Z_{t-2} + I_{t-3} Z_{t-3}$$

In order for this to work correctly, you need to make sure that the annual value is `NA` (missing value) for all periods except the last in each year. This won't happen if you pull in the data from, for instance, a RATS format file as an annual series, as it will just repeat the annual number in each quarter. If you need to patch the annual series, just do this:

```
set annual = %if(%period(t)==4,annual,%na)
```

In order for this to give (almost) exact answers, the initial variance of the states (`SX0`) needs to be "infinite." You may need to adjust the very large positive value given below (100000) if the scale of the problem requires it.

To adapt this to a quarterly to monthly problem, just change the 4's to 3's in the **DECLARE** instructions, remove the last row and column from `a` and `sw` and the last item from `c`.

This does an example included in the IMF description of benchmarking. The program below is available on "Procedures" page of the Estima web site as `DENTON.PRG`.

```
* DENTON.PRG Benchmarking Example
cal 1998 1 4
all 2000:4
data(unit=input) / quarter annual
 98.2 100.8 102.2 100.8 99.0 101.6 102.7
101.5 100.5 103.0 103.5 101.5
 . . . 4000 . . . 4161.4 . . . .
```

The dots are read as missing values in free format

```
print
dec rect a(4,4)
input a
 1 0 0 0
 1 0 0 0
 0 1 0 0
 0 0 1 0
dec symm sw(4,4)
input sw
 1
 0 0
 0 0 0
 0 0 0 0
dec frml[vec] c
frml c = ||quarter{0},quarter{1},$
           quarter{2},quarter{3}||
dec frml[vec] y
frml y = ||annual||
dec symm sx0(4,4)
comp sx0=%mscalar(100000.0)
dlm(type=smooth,a=a,c=c,y=y,sw=sw,sx0=sx0) $
   1998:1 2000:4 states
set z = states(t)(1)
set x = z*quarter
```

On-line Software Registration

If you have recently purchased one of our products, or would like to provide us with a change of address, we invite you to use our convenient new On-line Registration Form. This will ensure that you continue to receive future issues of this newsletter, new product announcements, and other information.

Registration is particularly important if your software was ordered through a reseller, or a university or corporate purchasing department—as we may have no idea who the actual end users are in such cases.

On-line registration is easy and fast, and is more accurate than the old method of using business reply cards (those hand-written responses are sometimes a little hard for us to read!). Just visit our website at:

http://www.estima.com

and click on the "Register Software/Change Address" button on the menu at the left of the screen. Please be sure to include your e-mail address in case we have any questions about your registration. And, if you are registering a software product, please be sure to include your serial number (it is printed on the CD envelope or on the diskette label).

Note that we *never* sell or distribute our customer database to anyone, so the information you supply will only be used by Estima, and only to send you issues of the *RATSletter* and other occasional mailings announcing new products.

RATS Version 5.02, *continued from p. 1*

```
dec rect shrink(n,n)
dec vector target(n)
compute shrink=%mscalar(la),target=(1-la)*mu
```

The **DLM** command to estimate this would include the options A=SHRINK,Z=TARGET.

If LA is a parameter being estimated, these would have to be replaced with FRML's:

```
dec frml[rect] shrink
dec rect sh(n,n)
dec frml[vect] target
frml shrink = (sh=%mscalar(la))
frml target = (1-la)*mu
```

Note that exogenous variables can be incorporated into the measurement equation by simply adjusting the Y formula to account for the shift.

New Functions

%(expression)—can be used to group sub-calculations where a comma would be interpreted as an argument separator. For instance, %if(x>=0,%(y=sqrt(x),x),y=%na) returns *x* if *x* is non-negative while setting *y* to the square root of *x*, and, if x is negative, returns a missing value, while also setting *y* to the missing value. In general, the "function" %(expr1,expr2,…,exprn) returns the value of the final expression in the group.

%TRADEDAY(t,d)—returns the number of occurrences of weekday *d* (Monday=1,Tuesday=2,…Sunday=7) in period *t*. For instance, January 2002 started on a Monday, and so has five of each of Monday through Wednesday, and four each of the other days. With a monthly **CALENDAR** in effect, %TRADEDAY(2002:1,1) returns the value 5, while %TRADEDAY(2002:1,5) returns the value 4.

%DAYCOUNT(t)—returns the number of days in entry *t*. For a monthly **CALENDAR**, %DAYCOUNT(2002:11) is 30, while if it's quarterly, %DAYCOUNT(2002:2) is 91.

%LNLOGISTIC(x)—returns $\log\left(\exp(x)/\left(1+\exp(x)\right)\right)$, that is, the log of the logistic cdf. Aside from simplifying the use of the logistic, it also handles tail values of x better than writing the expression out.

%EIGDECOMP(S)—returns a 2-vector of RECTANGULAR arrays which provides an eigen decomposition of the symmetric array **S**. The first array returned is the *N*x1 matrix of eigenvalues, and the second is the *N*x*N* matrix of eigenvectors. The eigenvectors are in the columns, and are normalized to unit length.

%COMPRESS(A,v)—returns an array with the same number of columns as **A**, but with all rows removed for which the corresponding element of the vector **v** is zero. **A** and **v** must have the same number of rows.

%MODELEQN(M,n)—returns the equation at position *n* in model *M*.

New Econometrics Books

Analysis of Financial Time Series, *by Ruey Tsay*
List Price $89.95, Estima's Price $75.00

Tsay's book is based upon an MBA course taught by the author at the University of Chicago. It covers a wide range of topics, from basic Box-Jenkins modeling, through ARCH and its relatives, duration models, continuous time models, value at risk (VaR) calculations, and multivariate time series and volatility analysis. It includes RATS programs for ARCH, non-linear volatility models and duration models. (SCA is used for the basic time series methods.)

Econometric Analysis of Cross Section and Panel Data, *by Jeffrey Wooldridge*
List Price $74.95, Estima's Price $65.00

Wooldridge's book is intended as a second semester graduate text. It examines the special problems that the econometrician must face in applying linear regression, instrumental variables/GMM and SUR estimators to cross section and panel data. It then covers a wide range of non-linear models: probit, logit, censoring and sample selection, count data and duration models. This includes almost all techniques covered in chapter 14 of the *RATS User's Guide* plus many more.

We've listed below the other econometrics texts available from Estima. These prices are subject to change, in the event of price increases by the publishing companies, so check the web site when ordering. Note that there are *no* shipping charges for domestic customers. International customers do need to pay for shipping, unless the textbooks are ordered along with RATS or a RATS update, in which case we generally do not add additional shipping charges.

Title/Author	List Price	Our Price
Applied Econometric Time Series	$86.95	$80.00
RATS Handbook for A.E.T.S.	$36.95	$34.50
both by Walter Enders		
Econometrics	$69.50	$55.00
Fumio Hayashi		
Time Series Analysis	$80.00	$65.00
James D. Hamilton		
Econometrics of Financial Markets	$70.00	$55.00
Campbell, Lo, and MacKinlay		

Important: New Credit Card Procedure

In an effort to combat credit card fraud, our credit card processing service now asks us to provide additional information with all credit card orders. On the back of your credit card, you will find a three or four digit security code (variously referred to as a CVV, CVC, or CIN code, depending on the credit card provider). This code is printed near the signature, usually following the regular card number.

If your card includes this security code, please be sure to supply it along with your credit card number, expiriation date, and billing address when placing an order. Thanks!

Panel Data, *continued from 1*

With either fixed or random effects, the panel data estimator can be written in the general form

$$(\mathbf{X}'\mathbf{W}\mathbf{X})^{-1}(\mathbf{X}'\mathbf{W}\mathbf{y})$$

Even though **W** is a big $(NT \times NT)$ matrix, it has a very simple structure for a balanced panel. This, however, breaks down when you have both time and individual effects in an unbalanced panel. To avoid the necessity of inverting directly a monster matrix, RATS relies upon a Kalman filtering procedure to take advantage of the sparsity of the matrix for which **W** is the inverse. For fixed effects, where the "variances" of the components are infinite, we use the variation on the Kalman filter described in Koopman's paper in the 1997 JASA (the full reference is in the description of the `KFEXACT` procedure in the next column). This is designed to handle analytically infinite initial variances.

This is simply a computational trick, which should produce exactly the same results as a brute force inversion or factorization of the covariance matrix. Where programs tend to differ is in the estimation of component variances for random effects estimation. There are three variances to be estimated. Three conditions are needed to tack those down. Several different "method of moments" conditions can be used, and, unfortunately, the results can differ rather dramatically depending upon which three are chosen. In our case, we changed the set being used between 5.00 and 5.01 because our original ones too often gave poor results.

Generalized Impulse Responses

Generalized Impulse Responses (Pesaran and Shin, (1998), "Generalized Impulse Response Analysis in Linear Multivariate Models", *Economics Letters*, 58, 17-29) are an attempt to avoid the difficulties of identifying orthogonal shocks in VAR models. This is not a particularly complicated idea; in fact, all these do is compute the shocks with each variable in turn being first in a Choleski order. These can be done quite easily using RATS; the following:

```
declare rect f(n,n)
ewise f(i,j)=sigma(i,j)/sqrt(sigma(j,j))
impulse(decomp=f) etc.
```

will give you the GIR.

However, while these can, quite easily, be done in RATS, we do not recommend them. While coming up with an orthogonalizing model can, in practice, be somewhat difficult, it *is* a necessary step. A set of responses to highly correlated shocks are almost impossible to interpret sensibly. For instance, you can't run an **ERRORS** instruction to assess the economic significance of the responses, since that requires an orthogonal decomposition of the covariance matrix.

The RATSletter

P.O. Box 1818
Evanston, IL 60204-1818
www.estima.com
Sales: 800–822-8038
Phone: 847–864–8772
Fax: 847–864–6221
estima@estima.com

New Procedures

We've posted several new procedures on the Estima web site (www.estima.com).

In addition to the procedures listed below, two example programs have been provided by Peter Pedroni for unit-root and cointegration testing with panel data. **`PANGROUP.PRG`** provides tests of specific hypotheses about cointegrating vectors in panels via group mean Fully Modified OLS techniques, while **`PANCOINT.PRG`** tests for the null of no cointegration in heterogeneous panels via both parametric and semi-parametric methods. There are some other recent additional contributions from our RATS users. Check out the procs directory on the web site.

`KERNREG.SRC` and **`LOWESS.SRC`** are procedures for flexible fits of the form $y = f(x)$. **`KERNREG`** does this by kernel regression (weighted sums of *y* values), while **`LOWESS`** uses locally weighted fits of first order polynomials. Both pick up some options from the instruction **`DENSITY`**. In particular, both allow you to provide the grid of *x* points at which you want to estimate the function, or to set the number of points and let the procedure set up the grid itself.

`KERNREG` also borrows the `TYPE` and `BANDWIDTH` options from **`DENSITY`**. With **`LOWESS`**, you indicate the stiffness of the function with the `FRACTION` option, which sets the fraction of data points which are to be included as neighbors for each test value of *x*.

`TSAYTEST.PRG` does a Tsay arranged regression test for the presence of threshold autoregression (TAR). This is a relatively simple procedure which demonstrates how the `ADD` option on **`KALMAN`** can be used to add data points to a regression out of sequence.

`KFEXACT.SRC` performs the Kalman filtering technique described in "Exact Initial Kalman Filtering and Smoothing for Nonstationary Time Series Models" by Siem Jan Koopman, JASA, December 1997. This is designed to deal with situations where there is no stationary initial distribution for the state vector. Koopman proposes this mainly to deal with an integrated series, particularly one with early missing values, but we have found the technique to be much more broadly applicable, as can be seen from our use of it in fixed effects estimation in panel data (other story on this page).

`REGTREE.SRC` performs a CART (Classification and Regression Trees) analysis. This is a nonparametric alternative to regression for analyzing the relationship between a dependent variable and a set of explanatory variables. CART works by splitting off the data into branches based upon the values of the independent variables. At each stage, it takes the branch which has the highest within group sum of squares for the dependent variable, and picks a split point which gives the greatest reduction in that sum.

The calculations required to create a CART are simple (nothing more than sums of squares), but tedious, as each value of each explanatory variable is examined at each stage.