The RATSletter

Volume 15, No. 2 **November, 2002**

Estima

Monte Carlo Integration, Revisited

In "Error Bands for Impulse Responses" in the September, 1999 issue of *Econometrica*, Christopher Sims and Tao Zha examine a large number of issues that are confronted when analyzing impulse responses for Vector Autoregressions. Among their recommendations are several which apply to all Monte Carlo analyses of impulse responses:

1. The responses often have a highly asymmetrical distribution. As a result, the use of one or two standard deviation bands can give a misleading impression about the shape. The authors recommend using fractiles instead, with .16 and .84 replacing the one s.d band, and .025 and .975 rather than two s.d.

2. A degrees of freedom correction should be included in the posterior density for covariance matrix of residuals.

We have posted to our web site (www.estima.com) three new or updated programs for Monte Carlo analysis of impulse responses. All of these incorporate the above changes. In addition, the array of graphs has been transposed from that in the previous MONTEVAR programs, so the responses of a single variable to the different shocks are in a single row, rather than a single column. That arrangement seems to be easier to read.

The second change above is quite simple. The first, however, requires more extensive changes. In order to compute fractiles, it's necessary to save all the draws, not just a few sums of sample values. Note that, in general, computing fractiles of a vector requires sorting it. (There are methods which are more efficient for computing a single specific fractile, but these aren't used in the RATS `%fractiles` function). And a sort will be required for each component of the desired impulse response function. For a six variable model with sixteen response steps, that's $6\times6\times16$ of them. If you're doing a large number of draws (10,000 or more), you might find that the time spent computing the fractiles is greater than that spent generating the draws.

The Sims and Zha article also describes a method for (correctly) computing error bands for overidentified structural VAR's. With an overidentified model, a draw from the posterior density cannot be generated by drawing an unrestricted covariance matrix, then mapping that to the restricted space. For the *Econometrica* article, the authors ended up using a very complex Markov Chain Monte Carlo algorithm. In our example program, we have, instead, implemented a more straightforward importance sampling technique, similar to that adopted in their earlier working papers. Because importance sampling is only minimally discussed in the RATS manual, we've included some additional background on page 2.

RATS 5.03

We've released another update to RATS—Version 5.03! This offers several useful new features, as described below.

If you have RATS Version 5.0, 5.01, 5.02 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. Macintosh and UNIX users can request a free 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.

You can also purchase a copy of the update on CD. The cost is just $15 for users who already have Version 5.0.

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

We've had a number of inquiries about purchasing "maintenance contracts" to avoid the small charges for the interim releases, and also to get the (future) update to version 6 on this year's budget. For a single user license, the cost is $150 to receive all 5.xx releases, plus 6 when it becomes available (probably late summer, 2003). The shipping charge is the same as our current charge for a RATS package ($0 for US, $22 for Canada and US Possessions and $50 elsewhere). Call for pricing on multi-user licenses.

A full list of changes is provided on the web site. The main changes with 5.03 are:

LINREG and **NLLS** now include an option for computing the optimal GMM weights when doing instrumental variables. Both will now automatically compute a test for overidentifying restrictions, and the computed weight matrix is available as `%WMATRIX`, which could then be passed through to another estimation.

MAXIMIZE now allows you to save the computed derivatives into a VECTOR of SERIES.

CVMODEL includes an option for choosing the precise form of the maximand. In addition to the current likelihood with the diagonal elements concentrated out, you can also choose the same with those integrated out in the posterior, or an unconcentrated likelihood, as recommended by Sims and Zha.

All the instructions which permit BFGS estimation (**MAXIMIZE**, **FIND**, **CVMODEL**, **DLM**) will allow you to input an initial estimate for the inverse Hessian using the HESSIAN option. This can help avoid problems with an inadequately formed inverse Hessian when you start out too close to the local optimum.

New Procedures

In addition to the new Monte Carlo programs, Wooldridge examples, and CATS example mentioned elsewhere, several other new procedures are now available for downloading from our web site. To access them, just go to:

www.estima.com

and click on the "Procs/Examples" button on the left.

We've recently redesigned the procedures section of the web site to make it more convenient. Rather than one (very) long page, there is now a "table of contents" page listing various categories of pages. Clicking on a category will take you to a page containing detailed information on the procedures in that category. We hope this will make it easier and faster for you to find and download the files you need.

Here are the files we've added recently:

RUNTEST.SRC performs a run test for independence for an input binary (0, 1) series. The probability of "success" can either be input or estimated.

SWAMY.SRC and **MEANGROUP.SRC** estimate panel data regressions using a random coefficients approach. The **Swamy** procedure is adapted from Example 14.3 in the *RATS User's Guide*. **MeanGroup** is similar, but takes scalar weighted averages of the regressions across individuals, rather than matrix weighted averages.

ROOTS.SRC includes the procedure **ComplexRoots** which computes the (complex) roots of an input polynomial.

WFRACTILES.SRC computes fractiles of a set of sample values with weights.

MATPEEK.SRC includes gtwo procedures: MatrixPeek and MatrixPoke which extract information from or put information into a rectangular array at locations specified in a list of (row,column) pairs.

FORCEFAC.SRC includes the procedure ForcedFactor which factors a covariance matrix with a (scale of a) specific vector either in a column of the factorization or a row of the inverse of it. This is useful when you want to analyze a specific linear combination of shocks without building a full structural model around them.

DLMGLS.SRC estimates a linear regression by GLS where the error process is described by a state space model like those used by the RATS instruction **DLM**.

Join us at ASSA Meetings

The ASSA convention returns to Washington, D.C. this January, and Estima will be there. The convention runs January 3rd–5th, 2003, at the Washington Convention Center. Our booth number is 407, and it appears to be located near the entrance, in the second aisle from the right.

We invite you to stop by and check out or software, database, and textbook products, ask questions, or just say hello!

Importance Sampling

The RATS 5 manual has a brief discussion of the use of importance sampling in Monte Carlo integration, but provides no examples. As one of our new examples relies crucially upon this technique, it makes sense to explain it more fully. For additional technical details, see Geweke, "Bayesian Inference in Econometric Models Using Monte Carlo Integration," *Econometrica*, November 1989, pp 1317-1339.

To review, importance sampling attacks an expectation over an unwieldy density function using

$$(1)\quad \begin{aligned} E_f(h(x)) &= \int h(x) f(x)dx \\ &= \int h(x)\big(f(x)/g(x)\big)g(x)dx \\ &= E_g\big(hf/g\big) \end{aligned}$$

where f and g are density functions, with g having convenient Monte Carlo properties.

The main problem is to choose a proper g. In most applications, it's crucial not to choose a g with tails which are too thin relative to f. If f and g are true density functions (that is, they integrate to 1), we can use

$$(2)\quad \hat{h} = (1/n)\sum h(x_i)\, f(x_i)/g(x_i)$$

to estimate the desired expectation, where the x_i are drawn independently from g. If $\int |h(x)| f(x)dx < \infty$, then $\int |h(x)\big(f(x)/g(x)\big)| g(x)dx < \infty$, so the strong law of large numbers applies to (2), and $\hat{h}$ will converge a.s. to $E_f(h(x))$. However, while the expectation may converge, if the variance of hf/g doesn't exist, the convergence may be extremely slow. If, on the other hand, that variance *does* exist, the Central Limit Theorem will apply, and we can expect convergence at the rate $n^{1/2}$.

To take a trivial example, where the properties can be determined analytically, suppose $h(x) = x$, and f is a Normal with mean zero and variance 4. Suppose we choose as g a standard Normal. Then

$$(3)\quad f/g = \frac{1}{2}\exp\left(+\frac{3}{8}x^2\right) \text{ and}$$

$$\int \big(h(x)f(x)/g(x)\big)^2 g(x)dx = \int \frac{x^2}{4}\exp\left(+\frac{1}{4}x^2\right)dx = \infty$$

Convergence of $\hat{h}$ with this choice of g will be painfully slow.

Suppose now that f is a Normal with mean zero and variance 1/4 and again we choose as g a standard Normal. Now

$$\int \big(h(x)f(x)/g(x)\big)^2 g(x)dx = \int \frac{4x^2}{\sqrt{2\pi}}\exp\left(-\frac{7}{2}x^2\right)dx = \frac{4}{7\sqrt{7}} \approx .216$$

For this particular h, not only does the importance sampling work, but, in fact, it works better than independent draws

(continued on page 4)

Arellano-Bond and other Panel Data

We've had a number of questions about how to implement the Arellano-Bond estimator in RATS (described in Wooldridge, *Econometric Analysis of Cross Section and Panel Data*, chapter 11). To summarize, this is applied in a panel data regression with fixed effects and a lagged dependent variable. If the data set is "small T, large N", a standard fixed effects estimator may be subject to a rather considerable bias. Instead, first differencing is used to eliminate the fixed effect, then lags (two and above) of the dependent variable are used as instruments. (Least squares applied to the first differenced regression would also be inconsistent). If you attempt to implement something like this directly in RATS, you run into the problem that, with T already being small, the use of a lagged dependent variable costs you one data point per individual, first differencing costs you another, then each lag (beyond two) used as instruments still another.

Instead of using the RATS "lags" for your instruments, you need to generate series which are lags if the data are available, but zeros if the lag would go out of the individual's data. The code below is a piece out of a RATS implementation of an example from the Wooldridge text. You can obtain the full example from the web site as described in the story in the next column. In this, `n` is the number of observations per individual, `lcrmrte` is log of the true dependent variable, and `clcrmrte` is the series of differences of `lcrmrte`.

```
dec vect[series] ablags((n-2)*(n-1)/2)
compute fill=1
do period=n,3,-1
  do lag=period-1,2,-1
     set ablags(fill)=$
       %if(%period(t)==period,lcrmrte{lag},0.0)
     compute fill=fill+1
  end do lag
end do period
instrument constant ablags
```

The next three instructions implement GMM with the optimal weight matrix, allowing for arbitrary serial correlation and **CDF** *tests the overidentifying restrictions.*

```
linreg(instruments) clcrmrte / resids
# constant clcrmrte{1}
mcov(lags=n-1) / resids
# constant ablags
linreg(instruments,wmatrix=inv(%cmom),$
        lags=n-1,robusterrors) clcrmrte
# constant clcrmrte{1}
cdf chisqr %uzwzu (n-2)*(n-1)/2+1-2
```

The sequence above can be replaced by a single instruction with RATS 5.03:

```
linreg(instruments,optimalweights,lags=n-1,$
       robusterrors) clcrmrte
# constant clcrmrte{1}
```

Economic Databases

A reminder that Estima offers two sources of economic databases—the OECD Main Economic Indicators database for international economic data, and Haver Analytics USECON, USNA, and US1 databases for U.S. data.

The OECD MEI provides data for all member countries and several of the most significant non-member countries. The data is supplied on CD ROM in the form of RATS format data files (one file per country). A "G7" subset of the database is also available, which includes data only for the original G7 countries. You can purchase a one-year subscription with monthly or quarterly updates, or you can order a single copy of the database.

The Haver Analytics databases provide detailed economic data for the U.S. USECON includes approximately 12,000 data series, including national accounts, prices, industrial production, money supply, etc. The USNA database adds an additional 20,000 series with complete national income and product accounts data from the Bureau of Economic Analysis. These series provide detailed information such as monthly personal consumption expenditures and personal income. The US1 database is a subset of the USECON database, containing approximately 750 of the most commonly-used data series.

The Haver data are supplied on CD ROM, both in RATS format, and in Haver's DLX (Data Link Express) format. The databases are offered as a one-year subscription. Commercial institutions will receive updates every month. Academic institutions have the option of getting monthly updates, quarterly updates, or of purchasing just a single copy of the database (the "Annual" subscription).

For more information or to place an order, please visit our website or contact us.

Wooldridge Examples

We've posted on the web site example programs and data sets for the Wooldridge book (*Econometric Analysis of Cross Section and Panel Data*, one of a number of texts which you can order from us). These provide fully worked examples of cross section and panel data techniques which are not given such a treatment in the RATS manuals. Among these are multinomial logit, ordered probit, duration models and Poisson regression.

To view or download the examples, go to www.estima.com and click on the "Procs/Examples" button to get to the Procedures and Examples page, where you will find a link to the Wooldridge examples.

CATS (Cointegration Analysis) News

We have a couple of pieces of news regarding the popular CATS in RATS cointegration analysis procedure.

First, the bad news. It has come to our attention that the critical values (90% quantiles) for the Lambda-max statistic reported by CATS are not correct. The L-max statistics themselves are correct, but the quantiles reported by CATS should not be used for making any inferences about the model. The Trace statistics and critical values (favored over L-max by the CATS authors), are correct.

If you do wish to make use of the Lamba-max statistic, you should use other critical values, such as those reported in Osterwald-Lenum (1992) or other simulated values. We should have a more detailed report on this up on our website by the time you read this.

Note that these critical values are only displayed for models with no exogenous or dummy variables, or if you explicitly specify the TABLES option.

Now some good news—many of you have been asking for a way to produce impulse responses based on a model estimated using CATS. Setting up the appropriate equations and getting the right coefficients out of CATS is tricky. So, we've written an example program that does most of the work for you. The program file is called CATSIRFS.PRG, and it is available for downloading from our website. The program does a call to CATS, then demonstrates the RATS instructions for setting up the appropriate equations (including identity equations), extracting the estimated coefficients from CATS and applying them to the equations, and then generating the responses. You could do variance decomposition easily as well, by adding the appropriate ERRORS command. To use this procedure, you also need to download a modified version of CATSMISC.SRC, that defines some required global variables.

RATS Internet Mailing List

The RATS internet mailing list suffered a couple of outages in recent months, but is now up and running again at full speed. For those of you who are not familiar with it, the "RATS List" is an e-mail forum that allows RATS users to share ideas, exchange sample programs and procedures, and ask questions. For details on the list and instructions on subscribing, see:

http://www.estima.com/maillist.htm

As always, our thanks to Rob Trevor for starting and maintaining this valuable resource for RATS users.

The RATSletter
© 2002 Estima
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

Importance Sampling, *continued from 2*

from the true *f* density. The standard error of the importance sampling estimate is $\sqrt{.216/n}$, while that for draws from *f* would be $\sqrt{.25/n}$. This result depends the shape of the *h* function- in this case, the importance function gives a lower variance by oversampling the values where *h* is larger. If *h* goes to zero in the tails, sampling from *g* will still work, but won't do better than draws from *f*. (The thin-tailed distribution works respectably only for such an *h*.)

The lesson to be learned from this is that, in practice, it's probably a good idea to be conservative in the choice of *g*. When in doubt, scale variances up or switch to fatter-tailed distributions or both.

All of the above was based upon *f* and *g* being true density functions. In reality, the integrating constants are either unknown or excessively complex. If, instead, we don't know f/g, just $w = f^*/g^* = cf/g$, where c is an unknown constant, then

$$(1/n)\sum h(x_i)w(x_i) = (1/n)\sum h(x_i)cf(x_i)/g(x_i) \to cE_f\left(h(x)\right)$$

$$(1/n)\sum w(x_i) = (1/n)\sum cf(x_i)/g(x_i) \to cE_f(1) = c$$

$$(1/n)\sum \left(h(x_i)w(x_i)\right)^2 = (1/n)\sum \left(h(x_i)cf(x_i)/g(x_i)\right)^2 \to c^2E_g\left(h(x)\,f(x)/g(x)\right)^2$$

Using the second of these to estimate *c* gives the key results

$$\hat{h} = \sum h(x_i)w(x_i) \Big/ \sum w(x_i)$$

$$s_{\hat{h}}^2 = \left[\sum \left(h(x_i)w(x_i)\right)^2 \Big/ \left(\sum w(x_i)\right)^2\right] - \hat{h}^2 \;/\; n$$

There's one additional numerical problem that needs to be avoided in implementing this. Particularly for large parameter spaces, the omitted integrating constants in the density functions can be huge. As a result, a direct calculation of w can produce machine overflows or underflows. (The typical computer can handle real numbers up to around 10^{300}). To avoid this, we would advise computing *w* by

$$\exp(\log f^*(x) - \log f^*_{\max} - \log g^*(x) + \log g^*_{\max})$$

where $f^*_{\max}$ and $g^*_{\max}$ are the maximum values taken by the two kernel functions.

Now all of the above shows how to compute the expectation of a measurable function of the random variable *x*. The fractiles needed for impulse responses don't fit directly into that framework. However, in general, the fractiles can be computed by sorting the generated values and locating the smallest value for which the cumulated normalized weights exceeds the requested fractile. The proof of this is in the Geweke article. The new procedure **WFRACTILES.SRC** on the web site computes these.