

## 2nd Edition E-Course Materials

The popular *State-Space and DSGE Models* and *Structural Breaks and Switching Models* e-courses have recently been updated to 2nd editions. The VAR course is also on its 2nd edition. The other subjects for e-courses are *ARCH/GARCH and Volatility Models*, *Panel Data*, and *Bayesian Econometrics*.

The e-courses go into considerably greater detail on specific subjects than we can provide in the main documentation. For instance, most of the technical discussion in this newsletter is a summary of longer sections in the *Structural Breaks and Switching Models* course. For more information, see

[https://estima.com/courses\\_completed.shtml](https://estima.com/courses_completed.shtml)

## Markov Switching GARCH Models

Hamilton and Susmel(1994), and independently, Cai(1994), proposed a Markov switching ARCH model as an alternative to a (standard) GARCH model. In these models, instead of the lagged variance term providing the strong connection for volatility from one period to the next, a Markov model governs switches between several variance regimes. Cai's variation on this is used in the SWARCH.RPF example from the *User's Guide*, while there's a replication example for the Hamilton-Susmel paper.

Given the dominance of the GARCH model over the ARCH in analyzing volatility, one might ask why those authors chose to do an MS-ARCH rather than MS-GARCH. There's a good practical reason for this—an exact analysis of an MS-GARCH model is nearly impossible because the lagged variance isn't observable and depends upon the entire prior history of the regimes up until that time. By contrast, the MS-ARCH uses only a finite number of lagged (observable) residuals.

To make MS-GARCH feasible, the likelihood must be approximated by using some method to summarize the regime history in a finite number of lags. There are two principal “filter” methods that have been proposed: Gray(1996) collapses it right away, so there is just a single lagged variance, while Dueker(1997) collapses it after a one period lag, so each regime at  $t-1$  has its own variance. Of the two, Dueker's is actually simpler to use in practice and more general and probably more accurate. Despite that, most other published work with MS-GARCH models has used Gray's filter.

(continued on page 3)

## How to Switch If You Must

We often have students who approach us with technical questions for the empirical chapter of a thesis where they have already largely committed themselves to doing a “Markov Switching Model.” And in far too many cases, the unfortunate answer has to be that a Markov Switching Model (MSM) really isn't what they need to analyze their data.

In fact, there are different types of models which allow for some type of “switching” behavior. The MSM is only one, and it is (probably by a wide margin) the most complicated and most likely to fail to produce reasonable results.

To be more specific, consider three models which are all based upon the following two “regime” linear model:

$$y_t = \begin{cases} X_t\beta_1 + u_t & \text{if } S_t = 1 \\ X_t\beta_2 + u_t & \text{if } S_t = 2 \end{cases}$$

The simplest possibility is the *structural break* model, which defines

$$S_t = \begin{cases} 1 & \text{if } t \leq T_0 \\ 2 & \text{if } t > T_0 \end{cases}$$

In most cases where the MSM is the wrong tool, this is what is *actually* needed. This is designed to look at a one-time change at a particular date (such as relaxation of currency controls, adoption of the Euro, etc.) Now, if you want to test for a change at a *specific* date, this is just a standard textbook “Chow test.” If you want to check more carefully whether a shift occurs at the date of interest (rather than some other date), then this becomes (for a linear model) an application of the @APBreakTest or @BaiPerron procedures.

Another model type has a *threshold break*

$$S_t = \begin{cases} 1 & \text{if } Z_{t-d} \leq \gamma \\ 2 & \text{if } Z_{t-d} > \gamma \end{cases}$$

where  $Z$  is an observable series, but the trigger value  $\gamma$  and possibly the delay  $d$  are unknown and must be estimated (typically by some form of grid search). This gives you different models for “high” and “low” values of the  $Z$  variable.

So how does an MSM compare with these? First, in the MSM  $S_t$  is unobservable—in fact, a more de-

(continued on page 4)

## New and Updated Programs

The files below contain example programs and (in most cases) data files replicating results from significant econometrics papers or demonstrating other useful techniques. These are available on our web site, and are included with version 9.2 of RATS. For many of these, there are fairly detailed descriptions as part of the RATS help. Note that some of these use features added with 9.2.

### balcilarguptamiller\_ee2015.zip

Balcilar, Gupta, Miller(2015), "Regime switching model of US crude oil and stock market prices: 1859 to 2013", *Energy Economics*, vol 49, 317-327. This is an example of a Markov Switching VECM (MS-VECM).

### balke\_restat2000.zip

Balke(2000), "Credit and Economic Activity: Credit Regimes and Nonlinear Propagation of Shocks," *Review of Economics and Statistics*, vol 82, 344-349.. This is an example of a threshold VAR (TVAR) and is covered in detail as part of the 2nd edition of the *Structural Breaks and Switching Models* e-course.

### bollerslev\_mikkelson\_joe\_1996.zip

Bollerslev and Mikkelson(1996), "Modeling and pricing long memory in stock market volatility", *Journal of Econometrics*, vol 73, pp 151-184. This demonstrates the FIEGARCH model (Fractionally Integrated Exponential garch).

### dueker\_jbes1997.zip

Dueker(1997), "Markov Switching in GARCH Processes and Mean-Reverting Stock-Market Volatility," *J of Business & Economic Statistics*, vol. 15, no 1, 26-34. This is one of two examples of Markov Switching GARCH models.

### ehrmann\_ellison\_valla\_el2003.zip

Ehrmann, Ellison, Valla (2003), "Regime-dependent impulse response functions in a Markov-switching vector autoregression model", *Economics Letters*, Vol. 78, pp.295-299. This is an example of a Markov Switching VAR and is explained in detail in the *Structural Breaks and Switching Models* course.

### grayjfe1996.zip

Gray(1996) "Modeling the conditional distribution of interest rates as a regime-switching process", *J. of Financial Economics* 42, 1996, pp 27-62. This is the second example of a Markov Switching GARCH model.

### hamilton\_susmel\_joe1994.zip

Hamilton and Susmel(1994), "Autoregressive Conditional Heteroskedasticity and Changes in Regime," *Journal of Econometrics*, vol 64, pp 307-333. This is an example of a Markov Switching ARCH (rather than GARCH) model.

### kilianvigfusson\_qe2011.zip

Kilian & Vigfusson(2011), "Are the responses of the U.S. economy asymmetric in energy price increases and decreases?", *Quantitative Economics*, vol. 2, no 3, 419-453. This does an analysis of an asymmetrical VAR, where positive and negative values enter the VAR with different coefficients. This uses bootstrapping to compute the non-linear impulse responses.

### matheson\_stavrev\_el2013.zip

Matheson and Stavrev(2013), "The Great Recession and the inflation puzzle," *Economics Letters*, vol. 120, no 3, pp 468-472. This is a bivariate "gap" model which allows for time-varying coefficients on the Phillips curve. This is covered in the 2nd edition of the *State-Space/DSGE* e-course.

---

## New and Updated Procedures

The following procedures are either new, or have had significant changes recently. There are quite a few others which have made more minor changes (corrections, or additional options). These are included in the 9.2 distribution, or can be download using the *Help-Update Procedures* operation. (Note that some require 9.2 and so will not be included in the updating process if you have an earlier version).

### @icss

This does the Iterated Cumulative Sums of Squares algorithm for searching for breaks in variance from Inclan and Tiao(1994). This has many new options to allow use of the identified breaks, and its use (and possible misuse) are discussed in the *Structural Breaks and Switching Models* course.

### @impactsignflip

This is a new procedure for fixing the impact signs in VAR models where a factoring procedure doesn't determine the signs (such as long- and short-run restrictions).

### @lsunit

This does the popular Lee-Strazicich unit root test allowing for trend breaks. We've reworked the critical values, which were done on a relatively coarse grid in the original paper, and added a **THIN** option to reduce the calculation time on larger datasets.

### @msregression

### @mssysregression

### @msvarsetup

These are the Markov Switching support procedures. They all have new procedures for more easily handling their **PARMSETS** and for doing guess values for the parameters. They've also been reworked to more easily handle time-varying transitions (TVTP).

## MS-GARCH, contd from page 1

We should note at this point, that several published papers (and many proposed projects from RATS users) have used or tried to use MS-GARCH when some other form of switching model was more appropriate. See the *How to Switch if You Must* article.

An interesting question is why the superior alternative was largely pushed aside. Probably the best explanation is that Dueker's *application* was not particularly convincing. Dueker analyzed daily S&P 500 returns. Despite trying four different combinations of switching parameters, none of his MS-GARCH models proved to be clearly better than an analogous non-switching GARCH.

By contrast, Gray worked with weekly US T-bill yields over a sample which included some fairly dramatic (and readily identifiable) changes, and his switching models fit substantially better than their non-switching counterparts.

In the 2nd edition of the *Structural Breaks and Switching Models* e-course, we demonstrate how to use both filters and apply both to the same model and data set (Gray's). Both filters give fairly similar results with two highly persistent regimes, but the Dueker filter produces an (approximated) log likelihood that is higher (by +10 on 1266 datapoints), most likely because the Dueker filter is more accurate (by carrying the history for one extra step).

However, it should also be noted that both filters with the same model and same data produce a second "mode" which has a much higher log likelihood (for Dueker's filter, 206 vs 171, with an even larger gap for Gray's). This other mode has regimes which aren't persistent and show very odd behavior (one with a lagged variance coefficient of 2.5). We had noted this in the replication program for Gray's paper that's been distributed with RATS for many years. Upon more careful analysis (using a particle filter), it turns out that the two filters fail badly to produce a good approximation to the unobservable lagged variance when the regimes aren't persistent, and that failure can produce a misleadingly high log likelihood.

If you're interested in this type of model, our recommendation is first, to make sure that it's really the type of model that will answer the questions you're asking, and if it is, use the Dueker filter and check the results carefully.

Cai, J. (1994): "A Markov Model of Switching-Regime ARCH," *Journal of Business and Economic Statistics*, 12(3), 309–316.

Other references can be found in the *New and Updated Programs* story

## State-Space/DSGE e-course

The *State-Space/DSGE* e-course covers most of Durbin and Koopman's *Time Series Analysis by State Space Methods, 2nd Edition*, with some additions from Harvey's *Forecasting, structural time series and the Kalman filter* and West and Harrison's *Bayesian Forecasting and Dynamic Models*. The textbook replication examples of all of those have many other examples of state-space models done using the RATS **DLM** instruction.

The second edition of this is greatly expanded compared with the first, with

- new sections on "gap" models with univariate models, bivariate models for gap plus Phillips curve and non-linear dynamic models
- greatly increased coverage of time-varying coefficients models
- a new section on Dynamic (Multiple) Factor Models
- a detailed discussion of Gibbs sampling methods
- a new section on particle filters
- a new section on Kalman filtering with constraints on the states

---

## Structural Breaks and Switching Models course

The *Structural Breaks and Switching Models* e-course treats a broad range of subjects, including tests for structural breaks and threshold effects, and estimation of threshold autoregression (TAR) and smooth transition (STAR) models. More than half the course is devoted to the topic of Markov Switching models, with applications to regressions, VAR, State-Space, and ARCH and GARCH models. It covers both maximum likelihood (EM where appropriate) and Bayesian estimation techniques.

The second edition adds over 100 pages, with new coverage of the ICSS test for variance breaks, the "fixed regressor bootstrap", increased coverage of computation of non-linear impulse response functions in various threshold models and a completely rewritten section on Threshold VAR's. We've reworked the various Markov switching support procedures, and the updated chapters on Markov Switching models and their examples have been revised to reflect that. In particular, there is now a (very) detailed description of the process of computing regime-specific impulse response functions with error bands.

## Switching Models, contd from page 1

scriptive title for these is Hidden (Regime) Markov Model. Second, and perhaps more important,  $S_t$  is (in most cases) uncontrollable. If more than one parameter is changing between the two regimes, it may be difficult to produce the desired separation between regimes. For instance, a common setup is to allow the  $\beta$ 's (or at least a subset) to change between regimes, with that being the intended interpretation, but to also allow the variances to change as well. Though the *intent* is that the variance is passively changing to follow the switches in the coefficients, far too often, with that model, the maximum likelihood estimates produce regimes which are best described as low- and high-*variance* and it's the coefficients that are just what they are in that regime breakdown. By contrast, with the other two types of regime models, you can observe the regimes (given the estimates) and in the case of the threshold model, can describe the conditions which separate the regimes.

So far this discussion has been about *linear* models, which are relatively simple to handle with structural or threshold breaks as they can be analyzed by running linear regressions across a grid of break values. Non-linear models are more difficult because they require repeatedly performing non-linear estimation. With modern computer speeds, this probably isn't so much an issue of calculation time (unless the data set is large or you're doing multiple breaks) but the difficulty in getting "NOPRINT" estimates of non-linear models to work reliably. GARCH models, for instance, are notoriously poorly-behaved when an outlier enters or drops out of an estimation window, and, in a threshold analysis, if an outlier drops out of one regime, it moves over into the other. Fortunately, if a non-linear model with a regime break is hard to fit, it will produce a lower log likelihood (or higher sum of squares for non-linear least squares) than properly converged estimates, and so will be much less likely to give the best break value for the model.

Note that models which have a recursively defined likelihood (GARCH, for instance), have to be estimated across the full sample using dummied variables to achieve the regime breakdown. The same is true with any model (even linear) that has coefficients that are common to the regimes.

---

The RATSletter

© 2018 Estima

1560 Sherman Ave, Suite 1029

Evanston, IL 60201 USA

[www.estima.com](http://www.estima.com)

[sales@estima.com](mailto:sales@estima.com)

874-864-8772

## Evaluation of GARCH Forecasts

A frequent request is for the calculation of the Root Mean Square Error (RMSE) or Mean Absolute Error (MAE) for a "GARCH model." The first question we need to ask is "errors in what?" GARCH models include both a mean model and a variance model, and sometimes the mean model is non-trivial (like a VAR).

For forecasting the data itself, the "GARCH" part of the model really doesn't matter unless it's a GARCH-M. (If it *is* GARCH-M, you need simulation methods to do the forecasts). If you use the MODEL option for a multivariate GARCH or EQUATION option for univariate, you can just use those on FORECAST or UFORECAST. However, it's important to note that these standard measures of forecasting accuracy (RMSE and MAE) treat all observations equally—they're really designed for homoscedastic error processes. Because of this, even if the GARCH process is correct, there is no reason to believe that the mean model estimated with GARCH errors would give better error statistics than simple least squares estimates.

Of greater interest would be the forecasts of the variance (volatility). The difficulty is that, unlike the data itself, the volatility is unobservable. An obvious choice for comparison is the squared residuals. However, as shown in Andersen and Bollerslev(1998), the squared residuals are a (very) weak proxy for the actual volatility. Now, in that paper, the authors use high-frequency data to provide an alternative measure of volatility to compare with the GARCH forecasts, and that type of data may not always be available.

If all you have is the squared residuals proxy, Patton(2011) looks at how various error statistics fare when used to evaluate volatility forecasts. He shows that the RMSE has the desired property of being smallest (theoretically, though noisily) for the correct model, but that the MAE does *not*—because the squared residual is non-negative and heavily skewed, an (incorrect) model which systematically underestimates the volatility will typically produce a lower MAE. Thus, despite the fact that these have been used in the literature, they aren't a useful tool for choosing a GARCH model.

Andersen, T.G., and Bollerslev, T., (1998), "Answering the Skeptics: Yes, Standard Volatility Models Do Provide Accurate Forecasts", *International Economic Review*, 39, 885-905.

Patton, A. J.(2011): "Volatility forecast comparison using imperfect volatility proxies," *Journal of Econometrics*, 160(1), 246-256.