

THE STATA NEWS

October/November/December 2011

Vol 26 No 4

Spotlight on state-space models: Easier than they look

State-space models are extremely flexible tools for both univariate and multivariate time-series analysis. Although they can appear intimidating at first, Stata commands like `ucm` and `dfactor` provide easy access to commonly used variants.

p. 2

Visit us at ASSA 2012

The Allied Social Science Associations (ASSA) will have their annual meeting in Chicago, IL, from January 6–8, and we will be there!

p. 3

Stata Conference

Call for presentations for the 2012 Stata Conference in San Diego, CA.

p. 4

Stata 12 now shipping

Still not using Stata 12? Here is just some of what you are missing...

p. 5

Check out stata.com on the go!

Some of the most popular pages on our website have been formatted for viewing on your mobile device.

p. 5

New from the Stata Bookstore

Using Stata for Principles of Econometrics, 4th Edition

p. 5

Public training course summaries

p. 6

Have you checked out the Stata blog lately?

p. 6

The Stata News:

Executive Editor:..... Karen Strope

Production Supervisor: ...Annette Fett

Learn Stata quickly via specialized training courses



Master the Stata features you need most by taking our new, focused training courses. Designed for rookies and established users alike, each of our new courses focuses on a narrow subset of Stata's capabilities, letting you learn what you need to know in as little time as possible. Whether you want to learn about multiple imputation, multilevel and mixed models, panel-data analysis, time-series methods, or survey data procedures, we have a course for you. Want to learn how to write your own estimation command that behaves just like Stata's official commands? We have a course for that too.

Each course lasts two full days and is held at a training facility in Washington, DC. Computers with Stata loaded are provided, though participants will want to bring a USB flash drive to save their work and datasets used during the course. The price to enroll is \$1,295, and each course includes a continental breakfast, lunch, and afternoon snack.

You can find brief course summaries on the back cover of this issue. Course details are available at stata.com/public-training.

Course	Dates
Handling Missing Data Using Multiple Imputation	April 4–5, 2012
Multilevel/Mixed Models Using Stata	February 9–10, 2012
Panel-Data Analysis Using Stata	April 18–19, 2012
Programming an Estimation Command in Stata	March 8–9, 2012
Survey Data Analysis Using Stata	May 30–31, 2012
Time-Series Analysis Using Stata	March 6–7, 2012

Don't forget about our general, two-day training course "Using Stata Effectively: Data Management, Analysis, and Graphics Fundamentals", which is held at various locations throughout the U.S. Visit stata.com/public-training for details.

Spotlight on state-space models: Easier than they look

State-space methods provide an integrated approach to modeling time-series data. Moreover, most widely used time-series models can be viewed in a state-space framework; and by recasting them as state-space models, we can expand them in ways that would otherwise be difficult if not impossible. For example, univariate ARIMA models, as implemented in Stata's `arima` command, are commonly used for forecasting and can be written in state-space form. The state-space framework handles multivariate data just as easily as univariate data, allowing us to fit multivariate generalizations of ARIMA models. Similarly, vector autoregressions (VARs) as fit by Stata's `var` command allow for multiple dependent variables, but they contain only autoregressive terms, not moving-average terms. The same state-space model that provides for multivariate ARIMA models allows us to fit vector autoregressive moving-average (VARMA) models that are like VARs but have moving-average terms as well.

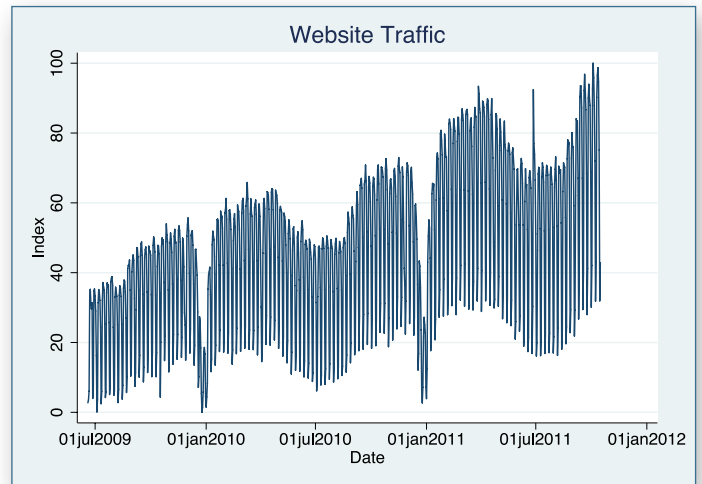
Because of their flexibility, state-space models can be difficult for newcomers to grasp. In some cases, the state-space representation is not even unique, adding to confusion; for example, there are multiple ways to express the widely used first-order moving-average model in state-space form.

Even for those who would rather not negotiate the details of state-space models, all is not lost. Some state-space models are so commonly used that Stata provides commands to make fitting those models easy; the commands hide the details of state-space representation so that you can focus on understanding and forecasting your time series. Here we focus on two models: unobserved-components models, implemented by the `ucm` command, and dynamic-factor models, implemented by the `dfactor` command.

Unobserved-components models

Most time series can be decomposed into as many as four distinct parts: a trend, a seasonal component, a cyclical component, and an idiosyncratic (random) component. Most series do not contain all four of these components, but the model is sufficiently flexible to allow all of them. Unobserved-components models allow us to estimate those four components, and they allow us to model the trend component in many different ways. Depending on the series, the trend might best be modeled as a deterministic process or a random walk with drift, or there might be no trend at all. Stata's `ucm` command makes fitting these models easy without having to invoke the full syntax of the `sspace` command.

Recently, I was asked to look at the following series, which represents the daily number of unique visitors to the Stata homepage (the numbers have been rescaled to range from 0 to 100):



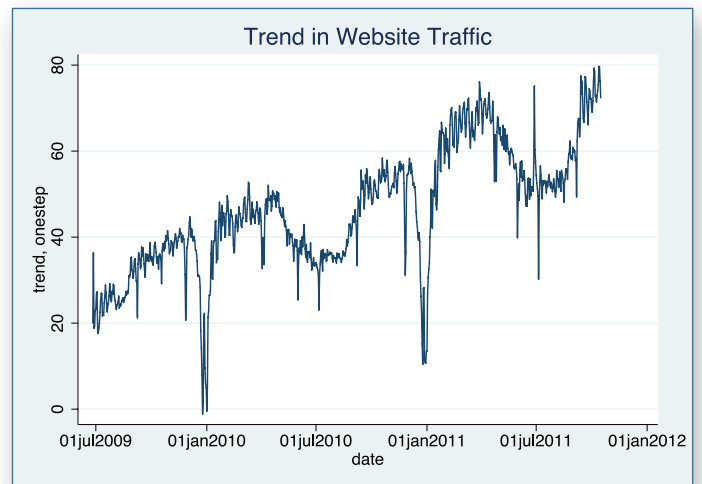
My boss wanted an answer quickly, but the weekly variation in traffic is hiding the underlying pattern. Stata's `ucm` command came to the rescue. I knew the period of the high-frequency fluctuations was 7 days, so I typed

```
ucm hits, seasonal(7) model(rwalk)
```

The option `model(rwalk)` tells UCM to model the trend component as a random walk. `ucm` allows you to select from 10 different types of trends or to not include a trend. I also tried modeling the trend using a random walk with drift, but both Akaike's and the Bayesian information criteria suggested the random walk without drift fits the data better. Once I had my model, I typed

```
predict trend, trend
```

and then used `tsline` to look at the underlying trend:



With the weekly fluctuations removed, I could more easily see the rises in traffic during the fall and spring semesters, the summer lulls, and the slow period around the winter holidays. Plotting the trend component also revealed other spikes in the data that were obscured by the weekly fluctuations.

I mentioned that unobserved-components models allow us to isolate a cyclical component as well, and you may be wondering why I used the `seasonal()` option here instead of the `cycle()` option of `ucom`. When you specify the `cycle()` option, `ucom` estimates the period of the cycle for you; in fact, you can have `ucom` detect up to three distinct cycles in your data. For example, you might be working with a time series from a scientific experiment and suspect a cycle but not know the precise frequency; `ucom` will estimate the period for you. In contrast, I knew that our website data fluctuates on a weekly basis and therefore did not have to estimate the frequency of the fluctuations. With the `seasonal(7)` option, `ucom` controls for the 7-day cycle but does not estimate its frequency directly.

Dynamic-factor models

Dynamic-factor models are flexible multivariate time-series models in which one or more unobserved factors influence observed dependent variables. The unobserved factors follow a vector autoregressive process, and the disturbances in the equations for the observed dependent variables may be autocorrelated. You can include exogenous variables in the equations for the unobserved factors and observed dependent variables as well. These models allow you to estimate the unobserved factors and make out-of-sample forecasts.

We have data on real consumption expenditures and disposable income, and we suspect that there is an underlying factor driving both of these (closely related) variables. We can use a dynamic-factor model to extract that signal:

```
dfactor (D.cons D.inc = , ar(1)) (f = , ar(1))
```

The syntax is not as complicated as it looks. First, consider this part of the command:

```
D.cons D.inc = , ar(1)
```

Because consumption and income trend upward over time, we are going to model them in first-differences; that explains the `D.` notation. The equals sign separates our dependent variables from the optional exogenous variables that can be included in the model; here we have no exogenous variables. After that, we specified the option `ar(1)` to indicate that we want each dependent variable to be modeled as a first-order autoregression. By default, lags of one variable do not appear in the

equations for the other dependent variables, but you can change that.

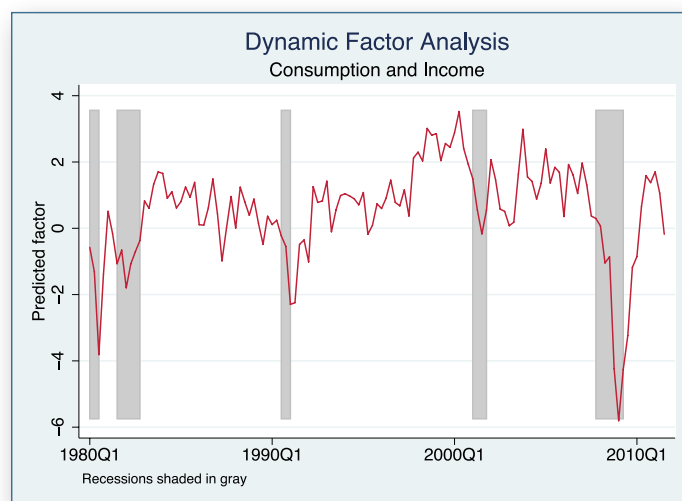
Now consider the second part of our `dfactor` syntax:

```
f = , ar(1)
```

This indicates that we have one unobserved factor, `f`, and that we are not going to include any exogenous variables in our model of it. The `ar(1)` option means that we are going to model `f` as a first-order autoregressive process.

Once we fit our model, we can then predict the unobserved factor:

```
predict factor, factors
```



(We used Kit Baum's SSC command `nbercycles` to produce the graph with recessionary periods shaded.) Because we modeled our dependent variables in first-differences, our predicted factor drives changes in those variables, not their levels directly. Notice how our predicted factor has downward spikes during recession periods.

Summary

State-space models are extremely flexible tools for both univariate and multivariate time-series analysis. Although they can appear intimidating at first, Stata commands like `ucom` and `dfactor` provide easy access to commonly used variants.

— Brian Poi
Senior Economist

Visit us at ASSA 2012

Chicago, IL, January 6–8, 2012

The Allied Social Science Associations (ASSA) will have their annual meeting in Chicago, IL, from January 6–8. For more information, go to aeaweb.org/Annual_Meeting.

Stata representatives, including Vince Wiggins, Vice President of Scientific Development, David M. Drukker, Director of Econometrics, and Brian P. Poi,

Senior Economist, will be available to answer your questions about all things Stata and demonstrate the new features in Stata 12. Stop by booth #803 to visit with the people who develop and support the software and to get 20% off your next purchase of Stata Press books and *Stata Journal* subscriptions.

We look forward to seeing you there!

STATA[®] Conference



Call for presentations

The 2012 Stata Conference will be held in San Diego, California.

The Stata Conference is enjoyable and rewarding for Stata users at all levels and from all disciplines. This year's program will include presentations by users, invited speakers, and StataCorp developers. In addition, the program will include the ever-popular "Wishes and grumbles" session in which users have an opportunity to share their comments and suggestions directly with developers from StataCorp.

All users are encouraged to submit abstracts for possible presentations. Presentations on any Stata-related topic will be considered, including (but not limited to) the following:

- new user-written commands, including commands for modeling and estimation, graphical analysis, data management, or reporting
- use or evaluation of existing Stata commands
- methods for teaching statistics with Stata or teaching the use of Stata
- case studies of Stata use in novel areas or applications
- surveys or critiques of Stata facilities in specific fields
- comparisons of Stata with other software or use of Stata together with other software

Each user presentation should be either 15 or 25 minutes long, and should be followed by 5 minutes for questions. Longer presentations will be considered at the discretion of the scientific committee.

Submission guidelines

Please submit an abstract of no more than 200 words (ASCII text, no math symbols) using the web submission form at repec.org/san/san12.php. All abstracts must be received by February 20, 2012. Please include a short, informative title and indicate whether you wish to be considered for a short (15-minute) or long (25-minute) presentation. In addition, if your presentation has multiple authors, please identify the presenter; the conference registration fee will be waived for the presenter.

If you would like to discuss an idea for a presentation or have questions about the program format, contact a member of the scientific committee.

Presenters will be asked to provide the organizers with electronic materials (a copy of the presentation and any programs or datasets, where applicable) so that the materials can be posted on the StataCorp website and in the Stata Users Group RePEc archive.

Scientific committee

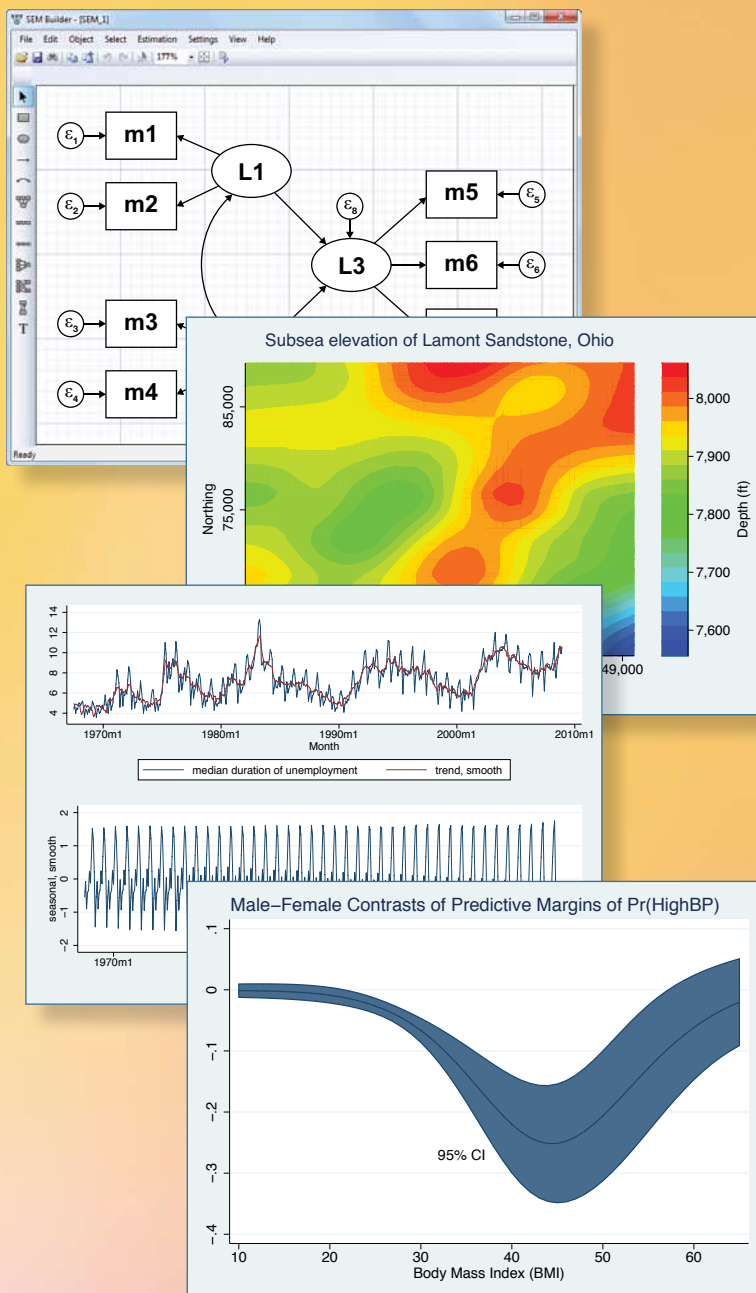
- A. Colin Cameron
University of California–Davis
- Xiao Chen
University of California–Los Angeles
- Phil Ender (chair)
University of California–Los Angeles
- Estie Hudes
University of California–San Francisco
- Michael Mitchell
US Department of Veterans Affairs

Dates	July 26–27, 2012
Venue	Manchester Grand Hyatt One Market Place San Diego, CA 92101
Cost	\$195 regular; \$75 student
Details	stata.com/sandiego12

Stata 12 now shipping

Still not using Stata 12? Here is just some of what you are missing:

- SEM (structural equation modeling)
- Excel® file import and export
- Mixed models with survey data
- Chained equations in MI
- Contour plots
- Margins plots
- Contrasts and pairwise comparisons
- Time-series filters: Hodrick–Prescott, Baxter–King, more
- Multivariate GARCH
- UCM (unobserved-components models)



Check out stata.com on the go!

Some of the most popular pages on our website have been formatted for viewing on your mobile device. To visit, just go to stata.com, and you will automatically be taken to our mobile home page.

Pages include:

- New in Stata 12
- Training
- NetCourses
- Short courses
- Order Stata (single-user)
- GradPlans



Of course, if you want to go to our full site at any time, you can click the link at the bottom of any page for complete access.

New from the Stata Bookstore

Using Stata for Principles of Econometrics, 4th Edition



Authors: Lee C. Adkins and R. Carter Hill
 Publisher: Wiley
 Copyright: 2011
 ISBN-13: 978-1-118-03208-4
 Pages: 611; paperback
 Price: \$68.00

Using Stata for Principles of Econometrics, Fourth Edition, by Lee C. Adkins and R. Carter Hill, is a companion to the introductory econometrics textbook *Principles of Econometrics, Fourth Edition*. Together, the two books provide a very good introduction to econometrics for undergraduate students and first-year graduate students.

The main textbook takes a learn-by-doing approach to econometric analysis, and this companion book illustrates the “doing” part using Stata. Adkins and Hill briefly show how to use Stata’s menu system and command line before delving into their many examples.

Using Stata for Principles of Econometrics, Fourth Edition shows how to use Stata to reproduce the examples in the main textbook and how to interpret the output. The current edition has been updated to include features introduced in Stata 11, such as the `margins` command to compute elasticities. Pairing this book with *Principles of Econometrics, Fourth Edition* will enable readers to not only learn econometrics but also gain the confidence needed to perform their own work using Stata.

You can find the table of contents and online ordering information at stata.com/bookstore/using-stata-for-principles-of-econometrics.

Public training course summaries

Handling Missing Data Using Multiple Imputation

This course will interactively cover all aspects of multiple-imputation (MI) analysis, including creation of MI data using the multivariate normal and chained-equations (or fully conditional specification) imputation methods, manipulation of MI data, and analysis of MI data. The course will provide exercises to reinforce the presented material.

Multilevel/Mixed Models Using Stata

This course is an introduction to using Stata to fit multilevel/mixed models. The course will be interactive, use real data, offer ample opportunity for specific research questions, and provide exercises to reinforce what you learn.

Panel-Data Analysis Using Stata

This course provides an introduction to the theory and practice of panel-data analysis. After introducing the fixed-effects and random-effects approaches to unobserved individual-level heterogeneity, the course covers linear models with exogenous covariates, linear models with endogenous variables, dynamic linear models, and some nonlinear models. An introduction to the generalized method of moments estimation technique is also included. Exercises will supplement the lectures and Stata examples.

Programming an Estimation Command in Stata

This course shows how to write an estimation command for Stata. No Stata or Mata programming experience is required. This course will provide an introduction to basic Stata do-file programming, basic and advanced ado-file programming, and an introduction to Mata, the byte-compiled matrix language that is part of Stata. Exercises will supplement the lectures and Stata examples.

Survey Data Analysis Using Stata

This course covers how to use Stata for survey data analysis assuming a fixed population. The course covers the sampling methods used to collect survey data and how they affect the estimation of totals, ratios, and regression coefficients as well as the three variance estimators implemented in Stata's survey estimation commands. Each topic will be illustrated with one or more examples using Stata.

Time-Series Analysis Using Stata

This course reviews methods for time-series analysis and shows how to perform the analysis using Stata. Exercises will supplement the lectures and Stata examples.

Using Stata Effectively: Data Management, Analysis, and Graphics Fundamentals

Aimed at both new Stata users and those who wish to learn techniques for efficient day-to-day use of Stata, this course enables you to use Stata in a reproducible manner, making collaborative changes and follow-up analyses much simpler.

We offer a 15% discount for group enrollments of three or more participants. Contact us at training@stata.com for details. For course details, or to enroll, visit stata.com/public-training.

Have you checked out the Stata blog lately?

Recent posts include:

- Good company (an analysis of software usage in published health services research)
- Multilevel random effects in `xtmixed` and `sem` — the long and wide of it
- Advanced Mata: Pointers
- Use `poisson` rather than `regress`; tell a friend

For helpful tips from top StataCorp staff and developers, visit blog.stata.com.

Contact us

979-696-4600

service@stata.com

Please include your Stata serial number with all correspondence.

979-696-4601 (fax)

stata.com

Find a Stata distributor near you
stata.com/worldwide



[facebook.com/StataCorp](https://www.facebook.com/StataCorp)



twitter.com/Stata



blog.stata.com



Copyright 2011 by StataCorp LP.

Serious software for serious researchers. Stata is a registered trademark of StataCorp LP. Serious software for serious researchers is a trademark of StataCorp LP.