

## Summer 2007 Session II Tutorial ECGA 8999

### Tutorial: Using Stata for Research in Economics, Sociology and Psychology

This tutorial meets Tuesday and Thursday 6:30-8pm starting July 5<sup>th</sup> in the Economics Conference Room, the July 5<sup>th</sup> class will meet 6-7pm and will serve as an orientation session. This course seeks to build research capacity in Fordham's graduate programs introducing students to latest statistical software and by encouraging the formation of "user groups" to support research in Economics and other graduate programs. Our goal this year is to establish a Stata users group at Fordham.

**Course Description:** The course introduces basic data handling and statistical analysis using Stata and Eviews. Students develop and present research applications starting with basic data entry and ending with final presentation of results, robustness checks and diagnostics in Journal style tables. Typically these applications replicate the result of a published article (using datasets available online) or with instructor approval data and analysis related to student research projects may also be presented. Topics covered vary with student interest, but will likely include panel and dynamic panel estimation, instrumental variables, duration or survival analysis, vector-error correction and auto regression (VEC and VAR) models, and limited dependent variable models. Students from Psychology, Sociology and Political Science are encouraged to attend, but may need to exercise more initiative in developing topics and data sets topics for analysis in their respective discipline.

**Prerequisites:** the course is open to graduate students who have completed two statistics courses, graduate or undergraduate. **Software:** We hope to have several copies of Stata 10 (released June 25<sup>th</sup>) installed in Economics and Psychology Department computers by July 10<sup>th</sup>. Students wishing to a Stata License using the "gradplan" should contact me at [mcleod@fordham.edu](mailto:mcleod@fordham.edu) or [dmcleod@pipeline.com](mailto:dmcleod@pipeline.com). Eviews is already available on many computers in IPED and in the E-530 Dealy.

**Short Description for Graduate Tutorial form:** You can download and print a [copy of the tutorial form](#) here. Use the 25 character title in blue above for the tutorial form and as wells as the following suggested short description: *Research applications using Stata or Eviews are developed by replicating the results of published articles or by students presenting their own research in standard journal table format complete with robustness and diagnostic tests.*

**Texts: Economics students:** Christopher F. Baum (2006) [An Introduction to Modern Econometrics Using Stata](#), Stata Press, 1-59718-013-0 (suggested readings [Chapter 1](#) for first lecture).

**Alternate Text:** Hamilton, Lawrence (2006) [Statistics with Stata](#) (updated for version 9) 0-495-10972-X, Stata Press, (make sure to get the 2006 edition).

**Before you come to class, get started with Stata using on line movie tutorials:** Even before you get Stata installed on your computer, you can review of the basics of data entry and manipulation using [online films and learning modules](#) prepared by the UCLA statistics department (if you have Stata installed, download the examples data using the instructions on this web page). You can start with this excellent [UCLA Stata](#) resources page. Those of you wish to listen in may be able to log in via Skype conference call contact me at [mcleod@fordham.edu](mailto:mcleod@fordham.edu) regarding access to [Skype](#) or just go [www.skype.com](http://www.skype.com) and sign up – its free with video and sound if you have a \$20 video cam. Check here Wednesday for the tutorial sign up form (or the course number for registration via Oasis).

Eviews: The IPED program uses Eviews and copies should be available on most Economics Department computers as well. Some examples Eviews 6.0 [Illustrated Graphics](#) Panel Unit Roots, etc.

## Stata-Eviews Users Summer Statistics Lecture and Tutorial Presentation Schedule

◀July 2007 ▶

Su Mo Tu We Th Fr Sa

<a href="#">1</a>	<a href="#">2</a>	<a href="#">3</a>	<a href="#">4</a>	<a href="#">5</a>	<a href="#">6</a>	<a href="#">7</a>
<a href="#">8</a>	<a href="#">9</a>	<a href="#">10</a>	<a href="#">11</a>	<a href="#">12</a>	<a href="#">13</a>	<a href="#">14</a>
<a href="#">15</a>	<a href="#">16</a>	<a href="#">17</a>	<a href="#">18</a>	<a href="#">19</a>	<a href="#">20</a>	<a href="#">21</a>
<a href="#">22</a>	<a href="#">23</a>	<a href="#">24</a>	<a href="#">25</a>	<a href="#">26</a>	<a href="#">27</a>	<a href="#">28</a>
<a href="#">29</a>	<a href="#">30</a>	<a href="#">31</a>	<a href="#">1</a>	<a href="#">2</a>	<a href="#">3</a>	<a href="#">4</a>

**July 10, 2007: Course Introduction:** Data entry using Stata using online data files, using the tutorial with text command files; importing data from spreadsheets into Stata w/o Stata Transfer. Checking the integrity of data before you do work with it (describe, summarize, graph box plot, histogram, etc.) **Readings: Baum Chapter 2.**

**July 12, 2007: Panel regression** as a general modeling framework: individual (fixed effects) and differencing: observations across entities (people or countries or localities) and over time: presenting data from regressions using excel spreadsheets; within and between estimators in Stata. Comparisons across disciplines: e.g. frailties vs. fixed effects.

**July 17, 2007: Introduction to Stata 10: Baum [Chapter 4 though section 4.4](#):** Presenting Regression results in Stata

**July 19 2007: Tutorial Presentation:** capital flows and growth panel regression using xtabond, xtabond2 and the new Stata 10 programs xtdp and xtdpdsys Elitza Mileva ( D. McLeod)

**July 24, 2007 Tutorial Presentation:** Rosendo Ramirez: Markups, Institutions and Growth: using Instrumental variable estimators in Stata: Baum Chapter 8.

**July 26, 2007: Tutorial Presentation:** Carolina Miccolis: Brazil Inflation, Growth and Money Demand: Times Series VAR and VEC models in Eviews and Stata?

**July 31<sup>st</sup>,2007: Tutorial Presentation:** Maria Davalos and Amy Rossnagel: Basic Logit and Probit Estimation in Sata 9.0 (10.0?)

**August 2, 2007: Tutorial Presentation: Wael Hibri and Daniel Scheer (TBA)**

**[Stata 10 Documentation](#)** (available on campus, in print form, but contact me for how to get access)



## Stata 10 Complete Documentation Set

Copyright 2007 ISBN

- 1-59718-027-0 Base Reference Manual (3 volumes), 1,886 pages
- 1-59718-030-0 Data Management Reference Manual, 548 pages
- 1-59718-034-3 Graphics Reference Manual, 618 pages
- 1-59718-031-9 Longitudinal/Panel Data Reference Manual, 475 pages
- 1-59718-037-8 Mata Reference Manual (2 volumes), 753 pages
- 1-59718-029-7 Multivariate Statistics Reference Manual, 621 pages
- 1-59718-032-7 Programming Reference Manual, 469 pages
- 1-59718-033-5 Survey Data Reference Manual, 171 pages
- 1-59718-028-9 Survival Analysis and Epidemiological Tables Reference Manual, 489 pages
- 1-59718-019-X Time-Series Reference Manual, 448 pages
- 1-59718-023-8 User's Guide, 356 pages
- 1-59718-020-3 Getting Started for Windows, 148 pages
- 1-59718-022-X Getting Started for Macintosh, 144 pages
- 1-59718-021-1 Getting Started for Unix, 152 pages
- 1-59718-038-6 Quick Reference and Index, 141 pages

[See a larger photo of the front cover of manuals](#)

[Overview of the Stata documentation](#)

Sample entries from the Base Reference Manual (pdf)

[clogit](#) — Conditional (fixed-effects) logistic regression

[mfx](#) — Obtain marginal effects or elasticities after estimation

[req3](#) — Three-stage estimation for systems of simultaneous equations

[regress](#) — Linear regression

[Download the datasets used in the manuals](#)

**Table of contents** *(Select a link to be directed to the table of contents for each manual.)*

**Base Reference Manual (3 volumes)**

[Volume 1 \(A–H\)](#) [Volume 2 \(I–P\)](#) [Volume 3 \(Q–Z\)](#)

[Data Management Reference Manual](#) [Graphics Reference Manual](#)

[Longitudinal/Panel-Data Reference Manual](#) [Mata Reference Manual](#)

[Multivariate Statistics Reference Manual](#) [Programming Reference Manual](#)

[Survey Data Reference Manual](#) [Survival Analysis and Epidemiological Tables Reference Manual](#)

[Time-Series Reference Manual](#) [User's Guide](#)

**Getting Started with Stata**

[for Windows](#)

[for Macintosh](#)

[for Unix](#)

[Quick Reference and Index](#)

## Data sets for practice and tutorials:

### [Stock and Watson, \*Introduction to Econometrics\*: data sets](#)

Each of these data sets is readable by Stata over the Web. You need only copy the line given below each dataset into your Stata command window or Stata do-file. After loading the data into Stata, use `save` to make a copy of the data on your own machine if you wish. The link from each dataset's name gives you the codebook of variable names and definitions. Please report any problems accessing these data to [baum](#).

- [CASCHOOL](#): N=420, panel data on test performance, school characteristics and student demographic backgrounds for California school districts, 1998-1999.
  - use `http://fmwww.bc.edu/ec-p/data/stockwatson/caschool`
- [CIG\\_CH10](#): N=96, Panel data, annual per capita cigarette sales for 48 states in packs per fiscal year in 1985 and 1995, only. use `http://fmwww.bc.edu/ec-p/data/stockwatson/cig_ch10`
- [CIG85\\_95](#): N=528, Panel data, annual per capita cigarette sales for 48 states in packs per fiscal year from 1985-1995. use `http://fmwww.bc.edu/ec-p/data/stockwatson/cig85_95`

[CPS\\_CH3](#): N=11130, Panel data: the Current Population Survey (CPS) provides data on labor force characteristics of the population, 1992-1998.

use `http://fmwww.bc.edu/ec-p/data/stockwatson/cps_ch3`

- [FATALITY](#): N=336, Panel data, number of state traffic fatalities for 48 states, 1982-1988. use `http://fmwww.bc.edu/ec-p/data/stockwatson/fatality`
- [HMDA\\_AER](#): N=2925, Cross-sectional data, mortgage applications made in 1990 in the greater Boston metropolitan area.
  - use `http://fmwww.bc.edu/ec-p/data/stockwatson/hmda_aer`
- [HMDA\\_SW](#): N=2380, Cross-sectional data, mortgage applications made in 1990 in the greater Boston metropolitan area using a subset of the original dataset.
  - use `http://fmwww.bc.edu/ec-p/data/stockwatson/hmda_sw`
- [JOURNALS](#): N=180, Cross-sectional data on 180 economics journals for the year 2000.
  - use `http://fmwww.bc.edu/ec-p/data/stockwatson/journals`
- [OJ](#): N=642, Time-series data on frozen and finished Florida orange juice prices, 1948-2001. use `http://fmwww.bc.edu/ec-p/data/stockwatson/oj`
- [MACRODAT](#): N=168, Time-series macroeconomic data. use `http://fmwww.bc.edu/ec-p/data/stockwatson/macrodat`
- [MCAS](#): N=220, Cross-sectional data, district-wide averages for Mass. public elementary school districts in 1998. use `http://fmwww.bc.edu/ec-p/data/stockwatson/mcas`
- [STAR\\_SW](#): N=11598, Subset of the variables in WEBSTAR.
  - use `http://fmwww.bc.edu/ec-p/data/stockkwatson/star_sw`
- [WEBSTAR](#): N=11598, Panel data on test scores, treatment groups, and student and teacher characteristics for four years (1985-86 to 1988-89).

use `http://fmwww.bc.edu/ec-p/data/stockwatson/webstar`

## Psychometrics Resource Page

[Rasch Analysis](#) [UCLA Course Notes](#) [Stata Routines](#) [Stata Support for Rasch](#)

[UCLA Course Example for Rasch Analysis](#)

[Gentle Guide to Stata](#) (by Alan Acock)

**Alan Acock** (Ph.D., Washington State University) is Professor and former Chair of Human Development and Family Sciences at Oregon State University. He has also taught at Louisiana State University, Virginia Tech, and the University of Southern California. Alan has published 4 books, 20 book chapters, and 120 articles. He is a Fellow of the National Council on Family Relations, a winner of the Reuben Hill Award, several awards for teaching, and his book on *Family Diversity and Well-Being* received the 1995 Choice Award for Outstanding Academic Book. Alan has held elected offices in the American Sociological Association and the National Council on Family Relations. His substantive research has been on the effects of family structure on the well-being of family members and on intergenerational relations. He has served on editorial boards of several substantive journals including the Journal of Marriage and Family. His methodological research has focused on structural equation modeling and missing values. He has one book on Stata and is currently writing a book with Peter Lachenbruch on Advanced Statistics Using Stata. His current statistical research is on applications of relative distributions to longitudinal interventions. He serves on the editorial board of the Journal of Structural Equation Modeling.

<http://www.hhs.oregonstate.edu/faculty-staff/userinfo.php?id=1>

### Journal Articles (sources for datasets models)

<http://www.psychometrika.org/journal/published/MostViewed2006.html>

<http://www.psychometrika.org/PDFs/ARTICLEstout2002.pdf>

<http://soldzresearch.com/statisticsresources.htm#Psychometrics>

<http://en.wikipedia.org/wiki/Psychometric>

<http://ssp.unl.edu/workshops.html>

<http://www.stata-press.com/data/acock.html>

<http://pareonline.net/>

<http://pareonline.net/genpare.asp?wh=0&abt=12>

## **Econometrics Course Tracks:**

Introduction to Econometrics using Stata:

- Entering and transforming data
- Graphical Display of Data using Stata and Eviews
- Simple Multivariate Regression Analysis
- Tests for Heteroscedasticity and Auto-Correlation
- Correlation and Rank Correlation
- Analysis of Variance / Principal Components

Intermediate Econometrics using Stata, Eviews and Pcgive::

- Seemingly Unrelated Regressions
- Panel Regression Techniques
- Series and Panel Unit Root Tests
- VARs and Autoregressive Models for forecasting
- Instrumental Variables: 2SLS and 3SLS
- Introduction to Arch Models

Advanced Econometrics using Stata, Eviews and Pcgive:

- Tests for Cointegration: the Johansen and Juselius
- Vector Error Correction Models of Money Demand
- Advanced IV tests and modeling techniques
- Impulse Response and Variance Decomposition
- Dynamic Panel Models
- Arch and Garch estimation
- Limited Dependent Variable Models (Logit and Probit)
- Duration or Survival Models: the Cox Hazard Model

Lecture notes July 5th 2007:

Objectives:

1. Make the transition to Stata: form a Stata users groups and encourage Fordham GSAS and CIMS to support Stata at Fordham University for graduate research.
2. Provide an online tutorial in an area of interest, explaining in steps using [screenshots](#) and explaining step by step starting with data entry.
3. [Replicate](#) the results of a paper or your own research using Stata or Eviews: present results to class with robustness checks, diagnostics and specification tests.

Texts: See syllabus: <http://www.gdsnet.org/classes/EconometricsSummer2007.pdf>

Software: Use the [Stata Gradplan](#) for Fordham University.

Schedule and Credits: Tutorial vs. Course