Econometrics for Analytics – MSA 8200
Spring 2016
(CRN 17158/18715)

Instructor Information

- **Instructor:** Daniel Bauer

- **Office:** Room 1126, Robinson College of Business downtown (35 Broad Street)

- **Office Hours:** By appointment (just send me an email, I am usually in most days)

- **Phone:** 404-413-7490

- **Email:** dbauer@gsu.edu

Class Information

- **Location:** GSU Buckhead Center, Classroom 404

- **Time/Dates:** 05:30-09:45pm. Classes will be held on 3/2/2016 (Wed), 3/8/2016 (Tue), 3/22/2016 (Tue), 3/29/2016 (Tue), 4/5/2016 (Tue), 4/13/2016 (Wed), 4/19/2016 (Tue), 4/26/2016 (Tue).

- **Website:** *D2L Brightspace* – `https://gsu.view.usg.edu/`

- **Prerequisites:** MRM8000 and MSA8190, or consent of the instructor. MSA students should follow the guidelines at `http://robinson.gsu.edu/msa/`, MRM students should follow `http://robinson.gsu.edu/mrm/`. You will need to be familiar with statistical programming environments. We will rely on R for the coursework. `http://www.r-project.org`.

Catalog Description

This course introduces students to econometric methods used in business analytics with a focus on real-world applications and datasets. The course covers two primary topics: econometric methods for panel data including how to account for basic heterogeneity effects; the most important models used for the analysis of time series including estimation and inference methods for univariate and vector auto-regressive models. After discussing these models in the classical context, the course revisits them using Bayesian methods with a focus on issues of parameter and model uncertainty. The course closes with a discussion of state-space models and Kalman filtering.

Textbooks

We will use two textbooks and provided material [P]. The textbooks are:

[W] Wooldridge, J.M., 2010. Econometric Analysis of Cross Section and Panel Data, Second Edition. MIT Press.

[L] Lütkepohl, H., 2005. New Introduction to Multiple Time Series Analysis. Springer.

In addition, I will post class notes on *D2L Brightspace*. Nevertheless, it is strongly advised that you take notes during lecture as there may be ideas presented in the class which **are not** included in the notes. Computer programs used in class will also be posted on *D2L Brightspace*.

Course Learning Outcomes

At the end of this course students will be able to:

1. Devise and estimate statistical models given a problem and data, and use the models for analysis.

2. Differentiate between methods and models suitable for "big" datasets, particularly depending on the characteristics (big N vs. big K vs. big G/T).

3. Understand when and why certain methods fail (e.g., OLS in case of endogeneity) and be able to conceive alternatives (e.g., instrumental variable regression).

4. Work with panel data and estimate corresponding regression models, also accounting for heterogeneous effects.

5. Model and forecast univariate and basic multivariate time series. In particular, estimate time series models and make predictions.

6. Deal with situations where there are many explanatory variables, either by selecting/condensing information or by relying on regularized estimation approaches.

Methods of Instruction

The material will be presented in lecture form. As a general approach, I first discuss models and methods conceptually (on the whiteboard!), and I then provide and discuss a variety of example problems/programs that illustrate the concepts. Depending on the subject, the theoretical and practical illustration may be in sequence or in parallel.

Attendance Policy

Attendance is not formally taken. However, it is strongly suggested that students do not miss class as most students will have difficulties completing the assignments without attending the lectures.

Exam Structure and Grading Criteria

There will be homework assignments to clarify and deepen concepts, an in-class mid-term examination, as well as a final project. You may discuss the assignments among each other, but every student has to write up the assignment on her/his own. Students copying from their classmates will receive a zero score. In addition, the student who let others copy from her/his assignment will receive a zero score. There are no exceptions to this rule!
The mid-term examination, homework, and the final project each as on third of the final grade. Make-up examinations are offered only under extraordinary circumstances. Students who miss examinations should contact me immediately. Grades will be awarded on a +/- basis, and the following guaranteed scale applies. Grades may be moved upward based on difficulty, but not downward

A+	A	A-	B+	B	B-	C+	C	C-	D+	D	F
98	95	90	85	80	75	70	65	60	55	50	< 50

Remarks

- Students exhibiting disruptive behavior, including talking, sleeping, talking on cell phones or disturbing other students will be asked to leave.

- Please advise the instructor if you have a documented disability that needs to be accommodated.

- As members of the academic community, students are expected to recognize and uphold standards of intellectual and academic integrity. See the Universitys policy on Academic Honesty (Section 409, `http://www2.gsu.edu/~wwwfhb/sec409.html`) for details.

- Accommodations for students with disabilities: Georgia State University complies with Section 504 of the Rehabilitation Act and the Americans with Disabilities Act. Students with disabilities who seek academic accommodations must first take appropriate documentation to the Office of Disability Services locate in Suite 230 of the New Student Center.

Detailed Outline:

- **Lecture 1: Introduction & Single Equation Models (Big N)**

  - What is "big data econometrics" / "statistics for analytics"?

    Big N (number of iid observations)

    vs. Big K (number of explanatory variables)

    vs. Big G/T (dimension of response variable/time-dimension)
  - Ordinary Least Squares Regression (OLS) and Endogeneity
  - Instrumental Variable (IV) Regression and Two-Stage Least Squares Regression (2SLS)
  - Some material beyond (difference-in-difference (Diff-in-Diff), specification tests)

- **Lecture 2: Systems of Equations (Big N)**

  - System OLS
  - Generalized Least Squares (GLS) Regression, Feasible GLS (FGLS)
  - Generalized Method of Moments (GMM) and System IV

- **Lecture 3: Nonlinear Models and Panel Data (Big N)**

  - Binary and Ordered Response Data
  - Pooled Ordinary Least Squares Regression (POLS)
  - Random vs. Fixed Effects

- **Lecture 4: Vector Autoregression Models (Big G/T)**

  - From Panels to Multivariate Time Series – what changes?
  - Vector Autoregression (VAR) Models

    Basics

    Estimation (Least Squares, Maximum Likelihood)

    Forcasting

- **Lecture 5: More on VARs (Big G/T)**

  - **MIDTERM** (first 90 minutes)
  - Causality and testing for it
  - Choosing models and testing their adequacy

- **Lecture 6: Multiple Time Series Beyond VAR (Big G/T)**

  - Cointegration & Vector Error Correction Models (VECMs)
  - Vector Autoregressive Moving Average Processes (VARMA)
  - State-Space Models

- **Lecture 7: Variable Selection, Dimension Reduction, and Regularization (Big K)**

  - Variable Selection: Mallow's Complexity Parameter (Mallow's CP), Akaike and Bayes Information Criterion (AIC/BIC)

      – Dimension Reduction: Principal Component Analysis

      – Regularized Regression: LASSO and Ridge Regression

- **Lecture 8: Catch-up and wrap-up, final discussion, Project discussion (Big $N, K, G/T$)**

      – **FINAL PROJECT** discussion

Date	Topic	Reading
2016-03-02	Introduction & Single Equation Models	[P] and [W] , Chap. 4-6
2016-03-08	Systems of equations	[W], Chap. 7-9
2016-03-22	Nonlinear Models and Panel Data	[W], Chap. 10 & 15-16
2016-03-29	VAR Models	[L] , Chap. 2-3
2016-04-05	**Midterm** (90 min) & More on VARs	[L], Chap. 4-5
2016-04-13	Multiple Time Series Beyond VAR	[L], Chap. 6-8, 11-12, 18
2016-04-19	Variable Selection, Dimension Reduction, and Regularization	[P]
2016-04-26	Catch-up and wrap-up, final discussion, **Project** discussion	

Note that this course syllabus provides a general plan for the course; deviations may be necessary.