University at Albany  
Rockefeller College of Public Affairs and Policy  
RPAD 705, Advanced Quantitative Analysis  
Fall 2016

**Time and Place:**
Mon 1:40PM-05:20PM  Husted 004 (we will only meet for 3 hours, we can talk about which 3 hours makes the most sense, in terms of schedules)  
Thursday 10:00am-11:30am  Husted 004

**Professor:**
Shawn D. Bushway Ph.D.  
324 Milne Hall  
Wk (518) 591-8738  
sbushway@albany.edu  
The best way to contact me is via email. To ensure that I see your email, write “PAD 705” in the subject line, and sign your full name.

**Teaching Assistant**
Cesar Renteria  
crenteria@albany.edu  
Cesar is a second year PhD student in PAD who is developing an expertise in econometrics. He has taken 688, 705, 704, and 725 and has taught statistics before.

**Office Hours:**
Bushway
I have found that students don’t make use of regularly scheduled office hours. However, I am in the office most days and I am willing to do whatever is necessary to facilitate appointments.

Renteria
Wednesday 4pm-7pm PhD Student Lounge, Third Floor.

**Course Goals:**
1) Students will have a strong intuitive understanding of multivariate regression, and its various uses, including statistical inference and statistical forecasting/prediction.  
2) Students will be able to perform sophisticated empirical analysis and interpret the results.  
3) Students will be able to read and critique empirical analyses used in academic publications.

**Prerequisites:**
I am assuming basic familiarity with statistics and basic regression at the level taught by PAD 505. Calculus is not required, but a conceptual understanding of calculus does make comprehension of the material easier. For PAD students, this course fulfills a core requirement, and, thus, your core GPA, which is evaluated as part of the Core Comprehensive review.
Expectations:
This is a 4 credit class. By university rules, that means that you are expected to spend 12 hours total on this class, 4 in class and 8 outside of class. I will not be surprised if you spend more than that, at least for some weeks. If you get above 15, let me or Cesar know. We might be able to help you study more efficiently.

I know of no one who has survived advance statistics without some help from a study group. Study together, but make sure you are doing your own work. In the end, your ability to understand the material will be evaluated on an individual basis.

The material in these books often includes equations. Many students are afraid of equations. Don’t be – equations are awesome. They are very precise and concise ways of communicating complicated concepts. Take your time, and don’t be surprised if it takes you an hour to read 5 pages. You are not stupid – you are normal. This is a different kind of reading, and it takes a little getting used to if you have never done it before. Don’t be afraid to start and stop, and bounce between books. Also, you are free to use outside resources – the web is full of material related to these topics, and you might be able to find things (including Wikipedia, believe it or not), which help you understand these ideas. Although we will necessarily have a theoretical focus, please always keep the applications in mind.

Book:
Required: We are trying a new book this year. It is a little less comprehensive, but it has an interesting approach, and is consistent with the goals of our program. We believe that it should be an excellent entry point into the material and a good reference for you throughout your career as a graduate student. I have given the names of both books to Mary Jane and the main bookstore.

Real Stats, Using Econometrics for Political Science and Public Policy.
Michael Bailey

Microeconometrics Using Stata, Revised Edition
A. Colin Cameron and Pravin K. Trivedi
Publisher: Stata Press
Copyright: 2010

Optional Texts
A Guide to Econometrics 6E Peter Kennedy
The edition doesn’t actually matter, and people have posted pdf versions on the web. Kennedy is to be used as a resource when you didn’t understand. Not all the explanations will makes sense (he occasionally goes at a higher level than we will cover in the class, but the genius of Kennedy is that everything is explained in three different ways, ranging from really basic to more complicated. The conceptual problems are good to

Basic Econometrics by Damodar Gujarati is an undergraduate econometrics text book that many people have found helpful. If you find yourself lost, head for Gujarati, and then work
A Full Course in Econometrics (on You Tube) by Ben Lambert
https://www.youtube.com/playlist?list=PLwJRxp3blEvZvQBTOMFRP_TDaSdly3gU

This is a nicely done lecture series on econometrics. Past students in 705 have found this very helpful. I do not lecture in class very much, so people who know they really like lectures might want to try these before class. They are also very good after you have been trying for a while – at that point his lectures can help lightbulbs go off.

A Gentle Introduction to Stata, Fourth or Fifth Edition
Alan Acock
If you are finding the Stata text too challenging, you can use this one as an intro. Not sure too many people will find it necessary. I also strongly encourage you to take Stephen Weinberg’s Stata class on Wed. mornings. Course Info: RPAD 688 Statistical Program Workshop
Meeting Info: W 10:00_AM-11:30_AM DR0014. (He requires the fifth edition). This is a required course for students in the Policy Concentration, and will have the biggest payoff if you take it early in your career.

Software:
The recommended software package for this course is Stata 13 (although older versions should be acceptable, and 14 will also work). You do not have to buy Stata -STATA is available on computers in the campus libraries and other laboratories throughout campus. When you are off campus, you can use the Virtual Information Commons to access Stata on your desktop. These resources should be utilized to complete your assignments.

If you choose to buy Stata, you can purchase it directly at the following URL:
http://www.stata.com/order/new/edu/gradplans/gp-campus.html

There are three versions of Stata:
1. Stata/SE or Stata/MP – If you plan to do a lot of statistical analysis for your dissertation, I would suggest buying one of these versions of Stata. You may also wish to buy the Base Reference Material and Stat Transfer.
2. Stata Intercooled (IC) – If you may do some analysis, but not with large datasets (ie. Not over 2047 variables), this version should suffice. Buy a temporary (one year) license if you never plan to do statistical analysis for your work.

Do not buy Small Stata. Some of the datasets used in this class will be too large for that version.

I will not teach STATA explicitly, but will use STATA in the class. Cesar (the TA) is a very good Stata programmer, and will be able to help you. In the end, you need to figure it out. The Cameron and Trivedi book should be a big help, and there are many resources available on the internet.
Here are two STATA websites that Cesar has found very helpful.

http://www.ats.ucla.edu/stat/AnnotatedOutput/ (basic stuff)
http://www.princeton.edu/~otorres/

In terms of STATA, let me state at the beginning that Cesar is a “much” better STATA programmer than I am. He should be your go to person for STATA questions.

**Classroom Approach**

This class uses a technique known as Team-Based Learning. Team-based Learning ([www.teambasedlearning.org](http://www.teambasedlearning.org)) is a type of active learning approach to classroom teaching that is part of the “flipped classroom” movement. The standard lecture class has the teacher present material in the class, and students do applications or learning exercises outside of class. In a flipped classroom, students do much of the concept acquisition outside of class through reading or other mixed media presentation, and then participate in applied learning activities in the classroom. Team-based learning is distinguished from other flipped classroom techniques by two features. First, team-based learning emphasized decision making during class time, in which students must use key course concepts to make decisions about real-life problems. Second, students spend all of their class time as members of permanent teams which facilitate decision making and ultimately, learning. Team work is done almost exclusively in the classroom, and part of the course grade comes from the team effort.

I first learned about Team-based Learning after arriving at UAlbany ten years ago. I have always done a lot of application but I struggled to motivate UAlbany students to apply concepts in the classroom. The technique has led to improved student outcomes (i.e. student learning) in my classrooms and I am a strong supporter of this approach both at UAlbany and at national conferences. For more information about my journey towards TBL adoption, see [http://www.itlal.org/index.php?q=node/287](http://www.itlal.org/index.php?q=node/287)

To see an academic article on Team Based Learning on which I am a co-author, please see the following website: [https://jstamatel.wordpress.com/courses/](https://jstamatel.wordpress.com/courses/)

**Academic Honesty**

The University’s standards for integrity are at the website below:
[http://www.albany.edu/graduatebulletin/requirements_degree.htm#standards_integrity](http://www.albany.edu/graduatebulletin/requirements_degree.htm#standards_integrity)

Avoid plagiarism by properly acknowledging material and ideas taken from other sources. The University of Albany Library offers a useful tutorial on plagiarism and how to avoid it: [http://library.albany.edu/usered/plagiarism/index.html](http://library.albany.edu/usered/plagiarism/index.html)

I take plagiarism very seriously and will report any incidents directly to the University.

**Grading Requirements and Procedures**

**RATS, Problem Sets, Papers and Exams:**

There will be 5 short, multiple choice Readiness Assessment Tests (RATs) given during the course. (The same RATs will be given to individuals, and then the Teams will take the same Test
again together. Both grades will count. There will be 6 problem sets. These will be graded (although not at the same level as an exam, and an answer key provided.) Students will be expected to study the answer key to complete the learning exercise. There will be a midterm and a final. Both will have an individual and a team component. These are graded separately. All students in the team get the team grade. The ultimate goal is to make sure that you can apply what you know to new situations.

The RATS and problem sets are tools that will help you be successful on the exams and papers. You will also read 5 mainstream PA papers in this class that use the tools we are learning. For one assignment, you will replicate that paper and produce a memo demonstrating the replication. We will give you the data, and we will be messing with the data so you will not be able to replicate exactly. The question, in part, will be to determine what is wrong with the data, and explain how it affected your replication. In the second paper, you will be asked to write a short paper extending the paper we replicated in the first analysis. Although you will be expected to write a short front end, the emphasis will be on the methods and results section. We will give you the correct data, and discuss ideas for extending the paper in class. In addition to a grade in the class, the students who earn a 95 or better on the final paper will collectively be given the following package of in kind services to help publish a paper (more than 1 publishable paper could arise from this process, but the resources are fixed).

1) A moderated meeting with Bushway to decide best way forward (either as a group, or as individuals)
2) 3 hours of Bushway consulting time to help with the paper
3) Blind reads/reviews by up to three faculty/colleagues who are working in this area
4) $200 towards professional copyediting services

**Grading Criteria:**
The grades will be determined by scores in three major performance areas:

**Individual Performance, Team Performance and Team Contribution.**

**Grade Weights and Percentages**

<table>
<thead>
<tr>
<th>Grade Weights within Area</th>
<th>Of Total</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1. Individual performance</strong></td>
<td>70%</td>
</tr>
<tr>
<td>Individual Readiness Assessment Tests</td>
<td>10</td>
</tr>
<tr>
<td>Problem Sets</td>
<td>10</td>
</tr>
<tr>
<td>Mid Term (in class)</td>
<td>20</td>
</tr>
<tr>
<td>Mid Term (replication memo paper)</td>
<td>20</td>
</tr>
<tr>
<td>Final (in class)</td>
<td>20</td>
</tr>
<tr>
<td>Final (paper)</td>
<td>20</td>
</tr>
<tr>
<td><strong>2. Team Performance</strong></td>
<td>20%</td>
</tr>
<tr>
<td>Team Readiness Assessment Tests</td>
<td>25</td>
</tr>
<tr>
<td>Midterm</td>
<td>35</td>
</tr>
<tr>
<td>Final Exam</td>
<td>40</td>
</tr>
</tbody>
</table>
### 3. **Team Contribution** (Evaluated by your teammates) 10%

**Evaluating Team Contribution:**
Each individual will rate the helpfulness all of the **other** members of their teams at the conclusion of the midterm and final exam. The midterm evaluation will be for practice and to help teammates adjust their behavior. The one after the final will be part of the final grade. Given that your teammates will be stuck with your during each and every class, they are in the best position to evaluate your contribution to the team.

Individual Team Contribution scores will be the average of the points they receive from the members of their team. Assuming arbitrarily that there five members in a team, an example of this procedure would be as follows. Each individual must assign a total of 40 points to the other four members in their team. Raters must differentiate between their teammates in their ratings and they can only give integer scores (This means that each rater would have to give at least one score of 11 or higher and at least one score of 9 or lower). As a result, Team Contribution scores will produce differences in grades only **within** teams. Consequently, team members can’t help everyone in their team get an A by giving them a high peer evaluation score. The only way for everyone in a team to earn an A is by doing an outstanding job on the individual and team exams and projects. These scores will be translated into a final participation score as follows:

Less than 20 D (60) (You have to be the world’s worst teammate to get less than 20 points in this system. I have only seen it happen once, and never in a PhD class.)

- 20-29 = C (75)
- 30 - 32 = B- (80)
- 33-38 = B (85)
- 39-42 = B + (88)
- 43-47 = A- (92)
- 48 or higher A (95)

**Three additional points:**
- a) Teams with more than/less than 5 teammates will use a slightly different, but analogous scale.
- b) You can collaborate to give each member of the team the same grade (40 points). This is fine, but you all have to agree ahead of time to get the same grade or it won’t work (there is no talking during the evaluations). If everyone collaborates, and everyone writes a statement indicating that they honestly believe that all the contributions were equal, a 40 will translate to a 90, rather than an 88.
- c) In addition to the point totals, you will be asked to provide feedback on your teammates strengths as well as provide constructive feedback on areas where the person could improve.

**Determination of Final Grades:**
The final grades will be determined as follows:
1) A raw total score will be computed for each student in each major performance area (In the
individual performance area, this will be a weighted combination of the sum of the individual
Readiness Assessment Test scores and the final exam score, in the team performance area, this
will be the sum of the scores on each of the graded team assignments and the Team
Contribution score will be the average of the peer evaluations received from the other members
of his or her team.)

2) Total scores will be computed by multiplying the raw scores in each area by the grade
“weight” (see above).

3) Course grades will be based on each individual’s standing in the overall distribution of total
individual scores within the class. (Note: When this procedure is followed: a) the actual impact of
any score on an individual student’s final grade depends on both his or her actual score and
also how high or low he or she scores relative to other members of the class and, b) the
conventional practice of 90% is an A, 80% is a B, etc. simply does not apply.) I will give all
A’s if everyone performs at an equally high level (this has never happened, but it could). In
general, I do not anticipate giving grades of less than B-, unless a student has clearly demonstrated
that they do not understand the material. This is a core class, and it is important that you take it
seriously.

Students with Disabilities
If you have a documented disability and wish to discuss academic accommodations with me,
please contact me. All reasonable efforts will be made to accommodate your needs.

Schedule – Please note this schedule is preliminary, and reading assignments may change. Please
check on Blackboard. Readings that are not from the required text will be available on
Blackboard.
<table>
<thead>
<tr>
<th>Date</th>
<th>Topic</th>
<th>Reading Assignment</th>
</tr>
</thead>
<tbody>
<tr>
<td>8/29/16</td>
<td>RAT 1 Introduction, Simple regression</td>
<td>Bailey Ch. 1, 2, 16, 3.1-3.3 (in that order)</td>
</tr>
<tr>
<td>9/1/16</td>
<td>In class Stata team exercises</td>
<td>Cameron &amp; Trivedi Ch 1 &amp; 2</td>
</tr>
<tr>
<td>9/5/16</td>
<td>Labor Day</td>
<td>No Class</td>
</tr>
<tr>
<td>9/8/16</td>
<td>In class Stata team exercises</td>
<td>Problem Set 1 Due</td>
</tr>
<tr>
<td>9/12/16</td>
<td>RAT 2 Bivariate Ordinary Least Squares</td>
<td>Bailey Ch. 3.4-3.8 &amp; 4 , Ch 14.1-14.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cameron &amp; Trivedi Ch. 4.</td>
</tr>
<tr>
<td>9/15</td>
<td>In Class OLS Exercises</td>
<td></td>
</tr>
<tr>
<td>9/19/16</td>
<td>In Class OLS Exercises continued</td>
<td>Problem Set 2 Due</td>
</tr>
<tr>
<td>9/22/16</td>
<td>RAT 3 Multivariate Regression</td>
<td>Bailey Ch. 5 &amp; 6</td>
</tr>
<tr>
<td>9/26/16</td>
<td>Multivariate Regression</td>
<td></td>
</tr>
<tr>
<td>9/29/16</td>
<td>Multivariate Regression, In class exercise</td>
<td></td>
</tr>
<tr>
<td>10/3/16</td>
<td>Rosh Hashanah – No class</td>
<td></td>
</tr>
<tr>
<td>10/6/16</td>
<td>RAT 4 Multivariate Regression Continued</td>
<td>Problem Set 3 Due</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Bailey Ch. 7, 14.3 to end</td>
</tr>
<tr>
<td>10/10/16</td>
<td>Multivariate Regression Continued</td>
<td>Academic Paper 1</td>
</tr>
<tr>
<td>10/13/16</td>
<td>Yom Kippur (Travel time) Optional Lab</td>
<td></td>
</tr>
<tr>
<td>10/17/16</td>
<td>Multivariate Regression Continued</td>
<td>Academic Paper 2 (Replication)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Problem Set 4 Due</td>
</tr>
<tr>
<td>Date</td>
<td>Event</td>
<td>Notes</td>
</tr>
<tr>
<td>--------</td>
<td>----------------------------------------------------------------------</td>
<td>-----------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>L7 10/20/16</td>
<td>Work on Replication</td>
<td></td>
</tr>
</tbody>
</table>
| C7 10/24/16 | Panel Data Replication memo due                                      | Bailey Ch. 8 (Fixed Effect)  
Cameron & Trivedi Ch. 8 (Optional) |
| L8 10/27/16 | Midterm During Lab                                                   |                                                                                               |
| C8 10/31/16 | Rat 4 Maximum Likelihood                                             | Eliason (1993) MLE: Logic and Practice Ch. 1-3,  
King (1999) Unifying Political Methodology: The Likelihood Theory of Statistical Inference Chapters 1-3 |
| L9 11/3/16 | Working on MLE                                                       | Bailey Ch.12;  
Cameron and Trivedi Ch. 14, 15;  
Andrich and Nelson (1989), Linear Probability, Logit and Probit Models Ch 1-3 |
| C9 11/7/16 | Logit Probit                                                         | Cameron Ch. 12;  
Cameron and Trivedi Ch. 14, 15;  
Andrich and Nelson (1989), Linear Probability, Logit and Probit Models Ch 1-3 |
| L10 11/10/16 | In Class Stata Exercise                                              | Academic Paper 4                                                                              |
| C10 11/14/16 | Count Models, Limited Dependent Variable                             | King (1999) Unifying Political Methodology: The Likelihood Theory of Statistical Inference Ch. 5.7-5.9, Ch. 9  
Cameron & Trivedi Ch. 16, 17 Problem Set 5 Due |
| L11 11/17/16 | In Class Stata Exercise                                              | Academic Paper 5                                                                              |
| C11 11/21/16 | Factor Analysis                                                      | Acock A Gentle Introduction Ch. 12  
| 11/24/16 | Thanksgiving No Lab                                                 |                                                                                               |
| C12 11/28/16 | Rat 5 Time Series                                                    | Bailey Ch 13, Pindyck and Rubinfeld Ch. 8, 15,16                                              |
| L12 12/1/16 | In Class Stata Exercise                                              | Academic Paper 5                                                                              |
| C13 12/5/16 | Bootstrapping, Errors                                               | Cameron & Trivedi Ch. 13  
Problem Set 6 Due |
| L13 12/8/16 | Work on Final Paper                                                  |                                                                                               |
| C14 12/12/16 | Final Paper                                                          | Final Papers due to group members and Bushway by 12/10                                       |
Bibliography


<table>
<thead>
<tr>
<th>Date</th>
<th>Event</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>12/19/16</td>
<td>Final Exam</td>
<td>3:30-5:30</td>
</tr>
<tr>
<td>3:30-5:30</td>
<td></td>
<td>at 5pm</td>
</tr>
</tbody>
</table>
STATA COMMANDMENTS
(BY AUDREY HICKERT, FOR THE LAB)

I
THOU SHALT ALWAYS USE A .DO FILE TO RUN AND SAVE COMMANDS

II
THOU SHALT USE THE MOST EFFICIENT COMMAND POSSIBLE (E.G., EGEN, LOOP, MACROS)

III
THOU SHALT JUDICIOUSLY ANNOTATE .DO FILES FOR LATER REFERENCE AND OTHERS' UNDERSTANDING

IV
THOU SHALT NOT OVERWRITE ORIGINAL VARIABLES

V
THOU SHALT ORGANIZE ALL RELATED PROJECT FILES IN A SINGLE DIRECTORY

VI
THOU SHALT LABEL FILES, VARIABLES, AND VALUES USING UNDERSTANDABLE NAMING CONVENTIONS FOR LATER REFERENCE AND OTHERS' UNDERSTANDING

VII
THOU SHALT SEPARATE .DO FILES FOR DATA CLEANING AND SUBSEQUENT ANALYSES ON FINAL ANALYTICAL FILE (CALLING .DO FILES WITHIN .DO FILES AS NEEDED)

VIII
THOU SHALT EXPLORE DATA DESCRIPTIVELY AND GRAPHICALLY (TO OBSERVE PATTERNS AND ANOMALIES) PRIOR TO ANALYSES

IX
THOU SHALT SEARCH FOR AND DOWNLOAD .ADO (USER WRITTEN SUCH AS OUTREG2 & TABOUT) FILES TO IMPROVE EFFICIENCY AND CAPACITY

X
THOU SHALT USE “HELP” IN COMMAND WINDOW TO SEE COMMAND OPTIONS, EXAMPLES, AND DESCRIPTIONS AND GOOGLE ERROR MESSAGES AND QUESTIONS TO REFERENCE HELP FROM THE STATA COMMUNITY