OVERVIEW: This course introduces students to the use of multiple regression analysis for analyzing data in the social sciences. The main goals of the course are for the students to be able to (a) understand the use of empirical analysis for addressing policy and management issues, (b) read and critique empirical analysis in academic and professional publications, and (c) perform such analysis and interpret the results themselves. Knowledge of statistics at the level of PAD 505 is assumed. Though calculus is not required, a conceptual understanding of calculus will make comprehension of the materials easier. For Public Administration Ph.D. students, this course is part of the Core and constitutes one component of the Core GPA, which is used to determine whether students may continue in the Program after their first year.

ADMISSION TO THE CLASS: All students must be enrolled in a Ph.D. program; Masters and Undergraduate students will not be admitted except in extraordinary circumstances. Students from the Public Administration Department are given first priority for slots in the class, which is limited to 20 students. All others will be admitted on a first come, first served basis, until the class maximum is reached.

READINGS: The required textbook for this course is *Econometric Models and Economic Forecasts (Fourth Edition)* by Pindyck and Rubinfeld, which is available at Mary Jane’s and the Bookstore. (I have asked the bookstore to make copies available in the Annex in the cafeteria.) Alternative presentations of the material covered in the course are contained in *Econometrics (Second Edition)* by Wonnacott and Wonnacott, *Basic Econometrics (Fourth Edition)* by Gujarati, and *Elements of Econometrics (Second Edition)* by Kmenta. All three are supposed to be on reserve in the library (though it appears that Wonnacott and Kmenta may be “lost”). The notation and presentation in class will follow Pindyck and Rubinfeld more closely, but most of the material is also contained in these other texts. The Wonnacott & Wonnacott and Gujarati texts are somewhat less difficult, while the Kmenta text assumes a higher level of mathematical proficiency than the others.
For those of you new to Stata, I also recommend purchasing *A Gentle Introduction to Stata* by Alan Acock, which is also available from Mary Jane, the campus bookstore, and the Stata order page (see below). Throughout the syllabus I will suggest some optional readings from *A Gentle Introduction*.

In addition, there are 26 (and possibly more) “Class Handouts” that are available from the course website:

http://www.albany.edu/faculty/kretheme/PAD705/overview.html

Go to the “Support Materials” section of the website to download the Handouts (available in PDF format). The Handouts address almost all of the major topics covered in the course. I will rely heavily on the examples presented in the Class Handouts during the lecture. *I strongly recommend reading the relevant Handouts before each class.*

Supplementary readings will be available through ERes.

The ERes system may be accessed from the library’s home page or from the course web site (see below). Find the readings for R. Karl Rethemeyer, select the PAD705 Fall 2008 option, and use the class password, pad705f08.

**SOFTWARE:** The recommended software package for this course is Stata 10.0. Stata is available for Windows, UNIX/Linux, and Macintosh. Stata may be purchased directly using the following URL or through the link on the course homepage:

http://www.stata.com/order/new/edu/gradplans/gp-campus.html

We have a so-called “Grad Plan3.” This means you order through the Web site but the software is distributed by me. Stata sends me a fax whenever an order has been completed online, giving me permission to release the software. There are several versions of Stata. Here are my recommendations:

1) If you plan to do a lot of statistical analysis for your dissertation, bite the bullet: Buy either Stata/SE or State/MP (multi-processor) and the Base Reference manuals (you’ll thank me later). Cost: $335 for SE or $695 for MP (2 processor) for the software; $179 for the Base Reference Manuals. You may also wish to purchase StatTransfer, a package that helps convert SPSS and SAS datasets into Stata format. Cost: $65.

2) If you think you might do some analysis for your dissertation but nothing that requires more than 2,047 variables, buy Stata Intercooled. You may also wish to buy the Base Reference manuals. Cost: $155 for the software; $179 for the Base Reference Manuals.

3) If this is the only class you plan to ever take and won’t do any statistical analysis for your work, buy the Stata Intercooled one year license. Cost: $95 for the software.

DO NOT buy Small Stata, which is the final option from the GradPlan page. Small Stata will not run some of the problem set datasets.

Stata is available in the Draper 09B, 015, and 023 computer labs for use on assignments, as well as most other student computer labs around the Uptown, Downtown, and East Campuses.

**ASSIGNMENTS:** Homework assignments must be handed in at the beginning of class on the day they are due. Assignments must be submitted *in class*. Late assignments must be submitted in my faculty box.
In Milne 103. Late assignments will assessed a 20% penalty, in part because the findings will be extensively discussed during the class in which they are due. For each week the assignments are late after the first week another 20% will be deducted. Students are strongly encouraged to work in small groups (2 - 4 people) but each student must write up his or her answers separately. Do not submit the same answer word for word. Answers that are identical to those of a study partner will receive no more than half credit on the first offense. The second offense will be considered academic misconduct and will be punished by other means (see below).

In addition to the regular assignments there will be a longer Empirical Exercise in the second half of the course. The Empirical Exercise is to be completed individually or with one other person.

There will be a 90 minute midterm exam and a 2 hour final exam. Both the midterm and the final will be open book / open notes.

GRADING: The final grade will consist of the homework assignments, the Empirical Exercise, the midterm, and the final, with the following weights:

<table>
<thead>
<tr>
<th>Assignment Type</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Homework Assignments</td>
<td>20%</td>
</tr>
<tr>
<td>Empirical Exercise</td>
<td>20%</td>
</tr>
<tr>
<td>Midterm Exam</td>
<td>25%</td>
</tr>
<tr>
<td>Final Exam</td>
<td>35%</td>
</tr>
</tbody>
</table>

I will also factor in class participation. Participation in lectures is at the student’s discretion. I will use attendance and participation as a “tie-breaker” if the grade falls at or near a “break-point” between grade levels. For instance, a person who has contributed regularly to class discussions and has a B+/A- average on the graded materials is more likely to receive an A- than a person who attends less regularly. As general policy, I will issue a grade of “C-”, “D”, or “F” if one’s average falls 1.5 or more standard deviations below the mean.

E-Mail communication
To reach me, use my personal e-mail address. However, for most communication please use the class LISTSERV. To subscribe, send an e-mail message to LISTSERV@LISTSERV.ALBANY.EDU with the line SUBSCRIBE PAD705-F07 YOUR_FIRST_NAME YOUR_LAST_NAME in the body of the message and nothing in the subject. You will be asked to confirm your membership in the list by a return message. To send a message to EVERYONE on the list, use the address

PAD705-F08@LISTSERV.ALBANY.EDU

Please register for this list as soon as possible and check your e-mail regularly for class news and information. If the class must be cancelled on short notice, the announcement will be made through the LISTSERV. Also use this LISTSERV for sharing common concerns and issues. Please do not use it for discussions or announcements that are not related to the class.

Time commitment for this course
This is a four-credit graduate course; it is also part of the Public Administration Ph.D. core. Hence you should plan on spending four to six hours per week in class and in the lab plus approximately six to eight hours per week doing the reading and preparing problem sets, worksheets, and cases. Students with strong prior background or experience in computing and/or statistics may spend less time than this. Students with little prior background may have to spend more time than this, especially in the first several weeks. If you discover that you are spending more than eight hours per week outside of class on this course, please let me know so that we can discuss it.
Plagiarism and cheating
Due to the intensive nature of this course, students are expected to form study groups and to work together on assignments. Learn by interacting with one another — support and help one another. However, (a) all students must submit an individually prepared copy of their homework (see above) and (b) some work such as the Empirical Exercise must be completed by the individual (or the individual and their partner) without collaboration with anyone else. As a policy for this course, plagiarism or cheating will result in a failing grade for the whole course. In addition, I will pursue further disciplinary action at the University level, including suspension and/or expulsion. For the purposes of this course, the following are taken as evidence of plagiarism or cheating:

- Material reproduced from another source without adequate citation.
- Identical answers being turned in by two or more students on a problem set, mid-term, final, or Empirical Exercise.
- Collaboration on the Empirical Exercise by two or more students without prior authorization.
- A pattern of unusually similar answers being turned in by two or more students on the Empirical Exercise, mid-term, or final.
- Written answers or solutions that a student cannot logically explain verbally.
- Other evidence of collaboration between students on an exam or the Empirical Exercise that was intended to reflect individual effort.

PLEASE NOTE: SEEKING PROBLEM SETS, ANSWERS TO PROBLEM SETS, PAST EXAMS, OR PAST EXAM ANSWERS FROM ANY PREVIOUS STUDENT IS PROHIBITED WITHOUT MY EXPRESSED, WRITTEN PERMISSION. I WILL TREAT SUCH BEHAVIOR AS SERIOUS ACADEMIC MISCONDUCT BY BOTH THE CURRENT AND PAST STUDENT.

Your work may be subject to computerized analysis to discover whether materials have been taken from on-line sources or to determine statistically whether answers are more similar than random chance would allow. Since this is such an important matter, if you have any questions about this course policy, you should ask me for any clarification that you may need.
## PAD 705 Research Methods II
### Detailed Listing of Assignments — Fall 2008

<table>
<thead>
<tr>
<th>Date</th>
<th>Topic</th>
<th>Readings Due</th>
<th>Written Work Due</th>
</tr>
</thead>
</table>
| 1. August 27 | Introductions & organization  
Comparing to experimental models  
Review of bivariate regression  
Multivariate regression | • Pindyck & Rubinfeld, Ch. 1, 2, 3.1-3.4  
• *Gujarati, p. 37-50; 119-133  
• *Acock, p. 1-75  
• H: An Introduction to Regression Analysis  
• H: Manual Calculation of Regression Parameters, t-Tests, and Confidence Intervals |                                                         |
| 2. September 2  
OR  
Another makeup date. | Problem Set #0 discussion  
Gauss-Markov Conditions  
Functional forms  
Dummy variables  
Interactions  
Goodness of fit | • Pindyck & Rubinfeld, Ch. 4.1-4.3; 4.5, 5.1-5.2, 7.3.3, Appendix 5.1  
• *Gujarati, p. 58-90; 107-113; 175-191; 297-323; 335-337  
• *Acock, p. 149-176; 211-224; 240-245  
• H: Functional Form & Goodness of Fit  
• H: Nonlinear Transformations  
• H: Omitted Variable Bias | Problem Set #0 (not graded) |
| 3. September 10 | Elasticities  
Standardized coefficients  
Hypothesis testing on multiple parameters  
Outliers & DFBETA  
Multicollinearity  
Stochastic explanatory variables, mismeasurement, and omitted variable bias (OVB) | • Pindyck & Rubinfeld, Ch. 4.4; 5.3; 5.5; 6.1; 7.1-7.4  
• *Gujarati, p. 257-273; 341-375; 387-504; 524-527  
• *Acock, p. 227-229  
• H: Hypothesis Testing on Multiple Parameters  
• H: Multicollinearity  
• H: Interpretation of Regression Coefficients When the Dependent Variable is in Logarithms  
• H: Standardized Coefficients  
• H: Outliers & DFBETA |                                                         |
<table>
<thead>
<tr>
<th>Date</th>
<th>Topic</th>
<th>Readings Due</th>
<th>Written Work Due</th>
</tr>
</thead>
</table>
| 4. September 17 | Problem set #1 discussion                  | • Pindyck & Rubinfeld, Ch. 5.5, 6.1-6.2  
• *Gujarati, p. 387-504; 510-513; 441-488  
• H: Heteroskedasticity  
• H: Heteroskedasticity, Robust Standard Errors, and Weighted Least Squares  
• H: Properties of OLS Estimators Under Heteroskedasticity                                                                 | Problem Set #1                        |
| 5. September 24 | Problem Set #2 discussion                  | • Pindyck & Rubinfeld, Ch. 6.2, 9.4  
• Gujarati, p 441-488; 636-652  
• H: Serial Correlation, the Durbin-Watson Statistic, and the Cochrane-Orcutt Procedure  
• H: Panel Data                                                                                                                                   | Problem Set #2                        |
| October 1     | Classes suspended                          | • Rosh Hashannah                                                                                                                                                    |                      |
| 6. October 8  | Panel data, continued                      | • Pindyck & Rubinfeld, Ch. 7.2.4; 9.4  
• *Gujarati, p. 636-652; 678-679  
• H: Instrumental Variables  
• H: Midterm Review 2003                                                                                                                        |                      |
| 7. October 15 | Problem set #3 discussion                  | • Pindyck & Rubinfeld, Ch. 12.1-12.3  
• *Gujarati, p. 717-729  
• H: Simultaneous Equations and Two-stage Least Squares                                                                                       | Problem Set #3                        |
<table>
<thead>
<tr>
<th>Date</th>
<th>Topic</th>
<th>Readings Due</th>
<th>Written Work Due</th>
</tr>
</thead>
<tbody>
<tr>
<td>TBD</td>
<td>OPTIONAL MID-TERM REVIEW</td>
<td>* All materials through October 15 lecture</td>
<td></td>
</tr>
</tbody>
</table>
| 8. October 22 | MID-TERM EXAM  
Simultaneous equations regression, continued – two-stage least squares | • Pindyck & Rubinfeld,; Ch. 12.4-12.5  
• *Gujarati, p. 735-756; 770-784 |                 |
| 9. October 29 | Angrist & Krueger discussion  
Introduction to maximum likelihood  
• Wise & Eichner manuscript (see ERes)  
• Pindyck & Rubinfeld, Appendix 2.2, Ch. 10.2-10.2.2, 11.1  
• *Gujarati, p. 114-117; 580-600; 608-609; 614-615  
• *Acock, p. 249-270  
• H: Maximum Likelihood Estimation  
• H: A Theoretical Discussion of Qualitative Dependent Variables  
• H: A Practical Introduction to Qualitative Dependent Variables  
• H: Diagnosing Output from Maximum Likelihood Estimations |                 |
| 10. November 5 | Problem set #4 discussion  
Regressions with qualitative dependent variables | • Pindyck & Rubinfeld, Appendix 2.2, Ch. 10.2-10.2.2, 11.1  
• *Gujarati, p. 114-117; 580-600; 608-609; 614-615 | Problem Set #4 |
| 11. November 12 | Advanced topics, qualitative dependent variables  
Empirical Exercise assigned | • Pindyck & Rubinfeld, Ch. 9.1-9.2  
• *Gujarati, p. 22-23; 656-664; 696-702  
• New handouts (to be announced in class) |                 |
<table>
<thead>
<tr>
<th>Date</th>
<th>Topic</th>
<th>Readings Due</th>
<th>Written Work Due</th>
</tr>
</thead>
<tbody>
<tr>
<td>November 26</td>
<td>Classes suspended</td>
<td>• <strong>THANKSGIVING BREAK</strong></td>
<td></td>
</tr>
<tr>
<td>December 8</td>
<td>Problem Set #6 discussion Optional Course Review</td>
<td>• Everything – but mostly since the mid-term</td>
<td>Problem Set #6 (turn in at review or in my faculty mail box, by e-mail, or by fax)</td>
</tr>
<tr>
<td>December 10</td>
<td>FINAL EXAM</td>
<td>• Everything – but mostly since the mid-term</td>
<td></td>
</tr>
</tbody>
</table>

* = Optional