RPUB/RPAD 705 – Research Methods II
Course Number: 5878
Spring 2003

Instructor: R. Karl Rethemeyer, Assistant Professor
Office: Milne 312A
Phone: (O) 442-5283
(H) 478-9599
E-mail: karetheme@albany.edu

Class meeting time and place; exam schedule
Monday, 5:45 PM to 8:50 PM in Richardson 290 – to be discussed at first class meeting.

Mid-term: March 24, 2003, to begin at 4:10 PM, in Draper 303. Ninety minutes.
Final: May 12, 2003, to begin at 5:45 PM (per University schedule). Two hours.
Room: Draper 303 (tentatively).

Note: Make-up exams will only be offered in the case of medical or family emergency or when mandated by University policy.

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 issues, (b) read and critique empirical analysis in academic and professional publications, and (c) perform such analysis and interpret the results themselves. The course covers the basic multiple regression framework, as well as extensions of that framework that become useful in actual applications of empirical policy analysis and research. A 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 covers materials that will be included on the comprehensive exam.

ADMISSION TO THE CLASS: All students must be enrolled in a Ph.D. program; Masters and Undergraduate students will not be admitted. 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 presentation of the material covered in the course is contained in Econometrics (Second Edition) by Wonnacott and Wonnacott, in Basic Econometrics (Fourth Edition) by Gujarati, and in 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 section references made in the syllabus may be out of sync with the newer version of Wonnacott & Wonnacott and Kmenta. 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. 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 Spring 2003 option, and use the class password, pad705s03.

**SOFTWARE:** The recommended software package for this course is Stata 8.0. Stata may be purchased directly using the following URL:

```
http://www.stata.com/info/order/new/edu/gradplans/gp3-order.html
```

We have a so-called “Grad Plan3.” This means you order through the Web site but the software is distributed by Professor Scott Barclay (POS). Stata sends Prof. Barclay a fax whenever an order has been completed online, giving him 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 Stata/SE and the Base Reference manuals (you’ll thank me later). Cost: $279 for the software; $129 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: $52.

2) If you think you might do some analysis for the dissertation but nothing very sophisticated, buy Stata Intercooled; you may also wish to buy the Base Reference manuals. Cost: $129 for the software; $129 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: $89 for the software.

Stata will *eventually* be available in the Draper 023 and 015 computer labs for use on assignments. However, the University has told us that Stata will not be available until February 17, at the earliest.

**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 at my office, and must be slid entirely under my door. Late assignments will not receive full credit, in part because the findings will be extensively discussed during the class in which they are due. Students are *strongly encouraged* to work in small groups (2 - 4 people) but each student must write up his or her answers separately. The assignments and associated data sets will be distributed through the course website:

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

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.*

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>Component</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 record attendance each evening and 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.

**E-Mail communication**
To reach me, use my personal e-mail address. To subscribe to this list, send an e-mail message to `LISTSERV@LISTSERV.ALBANY.EDU` with the line `SUBSCRIBE PAD705-S03 <FIRST NAME> <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** who is subscribed, use the address `PAD705-S03 @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 time than this on the 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 and (b) some work such as the Empirical Exercise must be completed by the individual 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 mid-term, final, Empirical Exercise.
- A pattern of unusually similar answers being turned in by two or more students on a 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.

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 2003 (Revised February 3, 2002)

<table>
<thead>
<tr>
<th>Date</th>
<th>Topic</th>
<th>Readings Due</th>
<th>Written Work Due</th>
</tr>
</thead>
</table>
| January 27 5:45 – 8:50 PM | 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 |                                               |
| February 3 4:00 – 7:00 PM | Gauss Markov Conditions  
Functional forms;  
Dummy variables  
Interactions  
Goodness of fit  
Elasticities | • 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 | Problem Set #0 (not graded) |
| February 10 5:00 – 8:00 PM | Hypothesis testing on multiple parameters  
Multicollinearity  
Stochastic explanatory variables, mismeasurement, and omitted variable bias (OVB)  
Heteroscedasticity | • 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 |                                               |
| February 17 | NO CLASS – PRESIDENT’S DAY |                                               |                                               |
| February 24 4:00 – 7:00 PM | Problem set #1 discussion  
Card & Krueger discussion  
Heteroskedasticity, continued  
Serial correlation | • Pindyck & Rubinfeld, Ch. 5.5, 6.1-6.2  
• *Gujarati, p. 387-504; 510-513; 441-488  
<p>| March 3 | NO CLASS – SPRING BREAK |                                               |                                               |</p>
<table>
<thead>
<tr>
<th>Date</th>
<th>Time</th>
<th>Event Description</th>
<th>Text References</th>
<th>Problem Set</th>
</tr>
</thead>
</table>
| March 10   | 4:00 – 7:00 PM | Serial correlation, continued Panel data: Fixed effects Random effects | • Pindyck & Rubinfeld, Ch. 6.2, 9.4  
• Gujarati, p 441-488; 636-652 | Problem Set #2 |
| March 17   | 5:00 – 8:00 PM | Problem set #2 & #3 discussion Panel data, continued Instrumental variables | • Pindyck & Rubinfeld, Ch. 7.2.4; 9.4  
• *Gujarati, p. 636-652; 678-679 | Problem Set #3 |
| March 20   | 5 – 7 PM     | OPTIONAL MID-TERM REVIEW                               | • All materials through March 17 lecture                                                              |             |
| March 24   | 4:00 – 7:00 PM | * MID-TERM EXAM * Introduction, simultaneous equations regression | • Pindyck & Rubinfeld, Ch. 12.1-12.3  
• *Gujarati, p. 717-729 |             |
| March 31   | 4:00 – 7:00 PM | Simultaneous equations regression, continued – two stage least squares Angrist & Krueger discussion | • Pindyck & Rubinfeld,; Ch. 12.4-12.5  
• *Gujarati, p. 735-756; 770-784  
| April 7    | 4:00 – 7:00 PM | Problem set #4 discussion Introduction to maximum likelihood Regressions with qualitative dependent variables | • Wise & Eichner manuscript (see ERes)  
• Wise & Rethemeyer manuscripts (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 | Problem Set #4 |
| April 14   | 4:00 – 7:00 PM | Regressions with qualitative dependent variables, continued Distributed lags, causality-intro | • Wise & Rethemeyer manuscripts (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 |             |
<table>
<thead>
<tr>
<th>Date</th>
<th>Time</th>
<th>Event</th>
<th>Reading</th>
<th>Topic</th>
</tr>
</thead>
<tbody>
<tr>
<td>April 21</td>
<td>4:00 – 7:00 PM</td>
<td>Problem set #5 discussion</td>
<td>• Pindyck &amp; Rubinfeld, Ch. 9.1-9.2&lt;br&gt;• *Gujarati, p. 22-23; 656-664; 696-702</td>
<td>Problem Set #5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Distributed lags, causality</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Empirical Exercise assigned</td>
<td></td>
<td></td>
</tr>
<tr>
<td>April 28</td>
<td>4:00 – 7:00 PM</td>
<td>Factor analysis</td>
<td>• Kim, J-O and Mueller, C. W. <em>Introduction to Factor Analysis</em></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Course wrap-up</td>
<td></td>
<td></td>
</tr>
<tr>
<td>May 9</td>
<td>4:00 – 5:30 PM</td>
<td>Optional Course Review</td>
<td>• Everything – but mostly since the mid-term</td>
<td>Problem Set #6 (may be submitted by e-mail or fax if you are unable to attend review)</td>
</tr>
<tr>
<td>Draper 303</td>
<td></td>
<td>Discussion of Problem Set #6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>May 12</td>
<td></td>
<td>Final Exam</td>
<td>• Everything – but mostly since the mid-term</td>
<td></td>
</tr>
</tbody>
</table>