Econometrics (Econ 308-04, #21765)
Department of Economics
The College of William & Mary
Spring 2023

Class meets Mondays & Wednesdays in Small Physics Lab, Rm. 235, 3:30-4:50 pm

Professor Peter Savelyev (pasavelyev@wm.edu)

About me My primary research interests are in the fields of health economics, applied econometrics, genoeconomics, and economics of human development. Prior to coming to William & Mary I worked at the University of Chicago and at Vanderbilt. See my Web page for more details: https://www.petersavelyev.com.

Office Hours Tuesdays 3:00-6:00 on for the periods of 2/5-3/11, 3/19-3/25, 4/2-5/5, or by appointment
Office Location: Chancellors Hall, Rm. 232

About this course This course is an introduction to econometrics, a collection of statistical methods for analyzing socio-economic data. Students will learn how to both analyze their own data and critically evaluate analysis done by others. Data analysis will include model specification, estimation and interpretation of model parameters, and hypothesis testing. The course will mix statistical intuition, the use of a powerful statistical package *Stata*, and mathematical/statistical derivations. The course emphasizes applications of econometric methods. Overall, this course provides students with essential statistical tools that are highly valued by employers and by graduate schools.

Prerequisites Econ 101, Econ 102, and Econ 307*
*Note: Students may use Buad 231, Math 106, Math 351, or Socl 353 in place of Econ 307 as a pre-requisite for ECON 308, but these courses do not count as credit hours toward the Economics major.

Stata Lab The Stata lab is organized by the economics department to serve multiple courses in economics that use Stata, not by myself for this specific course. However, TAs from the Stata Lab can serve as useful resource for this class. They can help with Stata issues and challenges with coding homework. They can also provide students with useful hints about homework.

These TAs are all experienced Stata users who have taken the course of econometrics already and have shown excellence in econometrics. Some of them have taken my class, some others have taken a similar class from another professor. Those TAs who have taken my class might be particularly helpful. All TAs have access to answers to my problem sets though the Stata Lab Blackboard cite. (Please remind them to look at the Stata Lab on Blackboard if they say they have no materials from me.)

Web page of the Stata lab:
https://www.wm.edu/as/economics/currentstudents/statalab/index.php

Reading Required reading is *Introductory Econometrics, A Modern Approach* by Jeffrey Wooldridge, 6th edition. Older or newer editions of the same textbook may work as long as students make sure that they do the right homework questions. Plus, lecture slides include material not covered by the textbook and may provide somewhat different interpretation of theory and results than in the textbook. Students should also know formulas and charts that I presented on whiteboard in class and Stata codes ("do-files") that we go over in class (all available on Blackboard).

Software This course relies heavily on Stata, a powerful and widely-used statistical package, arguably a #1 statistical software choice for most economists. Three typical ways to access Stata:

(1) On any W&M campus computer (find Stata software under "all programs/statistics" of the windows Start panel). A list of on-campus lab locations can be found here:
http://www.wm.edu/offices/it/services/computerlabs/configuration/index.php

(2) Using your own computer and remote access to Stata at W&M (from any place with an access to the Internet including off-campus). See below a paragraph called "How to run W&M Stata remotely from home."

(3) Stata can be installed to your personal computer if you choose to purchase it. The college is part of the Stata Grad Plan which offers discounted versions of Stata. More information on that can be found here: https://www.wm.edu/offices/it/services/software/licensedsoftware/mathstats/stata/index.php

Please do not leave your empirical work to the last moment, especially if you depend on a public computer or do not have experience working in Stata remotely!

Stata users enjoy a wide range of resources online in addition to excellent and user-friendly documentation that comes with Stata. You can Google an answer to almost any technical Stata question that you may have while solving the homework. One free introduction to Stata is called "A brief introduction to Stata with 50+ commands." https://pokrovka11.files.wordpress.com/2012/10/introduction_to_stata_with_50_basic_commands.pdf I will devote some time in the online class to learning Stata basics. That said, students are expected to get themselves familiar with the practical use of Stata by using material from class, from Stata help, and from web sources when needed. TAs at the Stata lab can be expected to be a useful resource.

Make sure you learn Stata well and add Stata skills to your CV. Not only you will be more successful in this course, you will also have better chances on the job market!

How to run W&M Stata remotely from Home

1. How to upload data to be used by Stata and how to start Stata remotely. (Note: remote Stata cannot use data files that are physically on your home computer. Instead, you need data files to be physically at W&M server.)
    1.1 Go to https://vdesktop.wm.edu/
    1.2. Select HTML access and go through W&M identification procedures (password etc.)
    1.3 Pick a computer lab from the list of available labs. If the chosen lab is busy, choose another lab.
    1.4 You will remotely see a screen of a W&M computer.
    1.5 Open a browser from the W&M virtual screen.
    1.6 Go to blackboard/econometrics/assignments and click on datasets_Stata_format.zip. This click will lead to downloading it and you will have an option to place data files to your folder at W&M (the H Drive, which is WM network storage).
    1.7 Go to the start menu and choose statistics/stata/stata. Stata window will open on your screen and you can start working. Open the data that you have put to your H-drive from Stata.

2. How to exchange files between work and home computers
    2.1 Do 1.1-1.5 as described above
    2.2 Connect to your favorite cloud storage to exchange files (e.g., Google Drive, One Drive, Box, Dropbox, iCloud). Now you can exchange files between cloud and W&M server. Therefore, you can move files from home to W&M and back though the cloud.

If these instructions do not lead to satisfactory remote connection, students should contact the Technology Support Center at support@wm.edu or by calling 221-HELP.

Some Useful Data Sources

Choosing, obtaining, and cleaning data takes remarkably long time, especially when you do so for the first time. Please do not leave your search of data to the last moment! Always choose public use data (unless you already have such data). You do not have the time to go through procedures to obtain the restricted data.

Here are some possible data sources in no specific order:

1. Integrated Public Use Microdata Series (IPUM): http://usa.ipums.org/usa/ Cleaned data here include the American Community Survey (ACS), Census data, Current Population Survey (CPS), and National Health Interview Survey (NHIS).
2. Multiple datasets from the NBER web page: http://www.nber.org/data/
3. NLSY data: https://www.nlsinfo.org/content/getting-started/accessing-data
4. MIDUS data http://midus.wisc.edu/data/index.php
5. WLS data: https://www.ssc.wisc.edu/wlsresearch/data/

6. Add Health data: https://addhealth.cpc.unc.edu/data/
7. Data that come with your textbook (see blackboard for this class or google it online). Be careful through that you do not use the same model that your textbook presents based on the same data. Such a paper will make no scientific contribution. You need to show your independent work and demonstrate some contribution.

This list above presents examples of datasets that students may choose. Students are expected to use their Internet search skills and find data source that matches their research interest. See handout 4 or Chapter 19 of the textbook about how to write an empirical paper and find data for it.

Group work Group work (or team work) is highly beneficial for learning since students learn from each other and get superior motivation. Moreover, group work creates skills that are essential for students' future careers: team work is a common work environment. I recommend students to form groups that consist of 2-4 people to work together on problem sets and the project. Please answer the questionnaire on group formation in the end of this syllabus and submit it as specified in the list of important deadlines (see below). One group member submits the questionnaire for the whole group.

Students are free to change the group, split the group, or start working alone at any time (after giving other group members a notice a reasonable time ahead of the next deadline). No approval is needed as long as your group does not exceed 4 students. Please view group work as your resource and learning opportunity, not a must. All members of the group will share the same scores for home assignments and their original research presentation. It is up to the group how they divide the work, however it is expected that contributions of all group members are substantial and comparable.

Grading Weights
    Homework (one work from each working group), 22%
    Midterm exam, 26%
    Group work in class, 2%
    Term paper (one paper from each working group), 15%
    Cumulative final exam, 35%

Midterm and Final Exams During exams students are allowed to use printed but not electronic books, lecture notes, own notes, printed Stata codes, and copies of homework and group work assignments. Computers, tablets, smart phones, and any other devices with Web access are not allowed. Calculators that have no Web access are allowed and might be useful.

The midterm exam takes the duration of one class. The cumulative final exam takes two hours.

If your score for the final exam is higher than your score for the midterm, then I will automatically replace your midterm exam score with the final score. For instance, if you get 85% for the midterm and 95% for the final, I will automatically replace 85% with 95%. However, if you get 80% for the final, I will make no such adjustments so that you

keep your midterm score of 85%. The aim of this policy is to encourage academic progress and give students a chance to improve their midterm scores.

Group work in class This work provides up to 2% of points for the semester-long effort. I can assign group work during any class, with no prior notification. The group work may be based on new material from the same class or any previous material. For the best results, students are encouraged to keep clean class notes, make sheets with formulas and definitions for quick reference, pay attention in class, and ask questions in class or during office hours if things are unclear. This policy encourages class participation and incentivizes students to pay attention in class and to stay on top of the material. Plus, students learn from each other while doing the class assignment.

Reward for active class participation Students who provide both useful and regular contributions to class discussions, ask good questions, and provide answers to questions that I ask the class during lectures will be rewarded by a higher grade in marginal cases (e.g, a B+ that is close enough to an A- will translate to A- for such student).

Homework Students will have weekly homework assignments, the majority of which involves computer work and interpretation of the results. Expect eight homework assignments (changes possible).

Please submit hard copies. Typing answers is encouraged. Handwritten formulas and charts are OK as long as your handwriting is clear. Unreadable answer will be considered as no answer.

If your homework requires Stata work, then print-outs of Stata's .do-files must be attached; points will be taken off if they are missing.

Unless announced otherwise, hard copies of homework assignments are due on Wednesday in class (see the list of important deadlines below). Groups have one week to complete homework (the homework is typically posted on Wednesday evening and announced by e-mail). Each group is allowed to submit one late homework to cover unexpected circumstances (skip the Wednesday deadline and submit it next Monday in class).

I will drop your lowest homework score to boost your average and total homework score.

Copying answers from other groups or any other source constitutes a honor code violation.

On each homework, please list the names of all group members who deserve credit for preparing that homework. To simplify grade recording, please use no name abbreviations, such as William of O. or Mary S.

Term Paper (Project) and its Abstract As stated in the course description, the goal of this course is for you to gain the ability to conduct and critique empirical work in economics and related fields. A testimony to whether you have achieved this goal is to implement an empirical project with your team and write a short paper. You will ask an

economic question, find relevant data, carry out empirical analysis, and write up a short term paper with other members of your group (I need one paper from each group).

You need to submit a short description (abstract) of your project by the deadline specified below in the list of important dates. Submit one printed abstract from each group. I will return it to you with my written comments. Please, limit the abstract to one paragraph. Motivate your project and mention which methods you plan to use, which data you plan to process, which results you expect, and which policy implications you may derive. I encourage students to discuss their topic with me any time when it is productive (before or after the abstract is due). The abstract will not be graded and is needed for the first feedback. Students are free to change their initial plan of work if needed upon a discussion with the professor.

Your project should contain the following parts:

1) Introduction: motivate your question of interest, shortly summarize which methods you used and which results you obtained

2) Data: Briefly describe your data. Define the outcome and main explanatory variables. Show a summary statistic table and, if needed, graphs summarizing important variables

3) Methodology: How do you answer the question raised in the introduction? Which econometrics method are you applying? What are the underlying assumptions? (For example, if you use a difference-in-differences estimator, clearly describe your control and treatment groups, and the assumptions which make such estimation valid.) Admit limitations of your methodology. For instance, unless you have data from a randomized experiment, a simple OLS model will likely fail to give you estimates that can be interpreted as causal under realistic assumptions. Estimates based on simple OLS are still fine for this course but you need to note the limitations and view you results as associations.

4) Results: Present your results in well-formatted tables and figures. Place notes behind tables and figures to make them self-explanatory. Every table and figure should have a clear title. Interpret your results. Mention limitations of your results related to limitations of methodology that you discussed above (for example, if you suspect your analysis suffers from omitted-variable bias but cannot do anything about it, acknowledge it, and, if possible, discuss the direction of the bias. Admitting limitations improves the quality of your project). Discuss both the economic and statistical significance.

5) Conclusions: Summarize what the paper is about and your findings. Discuss policy implications if any exist. Mention possible steps you may want to undertake in the future to move this research forward.

6) An appendix showing a do-file that produce results of this paper including data cleaning and results calculations.

The main text of the project report should be no more than10 pages double-spaced using 12-point font (figures, tables, and an appendix with a do-file and a log-file are excluded

from this restriction). Smaller fonts and different line spacing can be used for tables, figures, and the appendix.

Examples of projects in the past Please find below examples of paper titles that were submitted for this class (all topics originated from students). Note that titles are very specific, and so these papers deeply study a specific research question. Your group can pick any specific topic including topic related to examples listed below.

- Household Composition and Student Performance
- The Effect of a "Right-to-Counsel" Law on Eviction Rates in New York City
- The Relationship Between Global Development and Life Expectancy
- Cigarette Excise Taxation and Intrastate Smuggling in Virginia
- Predicting Starting Salaries of Lawyers Using Education-Related Factors
- Understanding a Negative Association Between Napping and Wages
- A Cross-Industry Analysis of the Gender Wage Gap: Comparing the Textile and Mining Industries of the United States

Avoid a typical mistake of picking a topic that is too broad for a paper. For instance title "The Effects of Education, Experience, and IQ on Wages, Health, and Happiness over the World" is way too broad for a paper. You will not be able to answer all these questions in enough detail in a single paper. In contrast, title "The Effect of Education on Mental Health in the US" is well-focused.

Important Dates
- 01/25: the first class
- 01/30: group formation questionnaire due in class
- 02/08, 02/15/, 02/22, 03/01, 04/05, 04/12, 04/19, 04/26: home assignments due in class (assignments will be posted on the Blackboard about a week ahead of the deadline)
- 03/08: midterm exam in class
- 03/11-03/19: spring break
- 03/22: title and abstract of the term paper due in class
- 05/03: the last class
- 05/04: pdf of the paper due by 10pm to a designated Blackboard folder.
- A two-hour cumulative final exam in the same room as the class. Note: this time and date is set by the college and subject to change by the college. The most updated official date of the exam can be checked here.
  - Exam time: Tuesday, May 9, 10:00 am to noon.
  - *Note:* officially, we have the room reserved from 9am to noon, but since the exam takes two hours only we can start at 10 am.

Expected topics by class number Numbers of chapters refer to Wooldridge's textbook, 6th ed., as the main source for this course. Other sources will supplement the textbook when preparing lectures but Wooldridge is the only required textbook. "HD#" refers to numbered handouts (which are lecture presentation slides) available on Blackboard. Topics include theory, Stata coding, and possible in-class work including exam preparation activities. Dates are expected ones and subject to changes.

1. Introduction (syllabus), class 1
2. Nature of Econometrics and Economic Data (Ch. 1, HD1), class 2
3. Multiple Linear Regression (MLR): Estimation (Ch. 2–3, HD 2); Introduction to Stata coding. OLS estimation and Monte-Carlo simulation in Stata, classes 3–7
4. MLR: Statistical Inference, related Stata options, and midterm preparation (Ch. 4, HD 3), classes 8–11

[Midterm exam during class 12 covering all topics so far]

5. Carrying out an Empirical Project (Ch. 19, HD4), class 13
6. MLR: Large Sample Properties (Ch. 5, HD5), class 14
7. MLR: Further Issues: Functional Form, Goodness of Fit and Model Selection, Prediction and Residual Analysis (Ch. 6, HD6), classes 15–16
8. MLR with Qualitative Information and the Linear Probability Model and related Stata options (Ch. 7, HD7), classes 17–18
9. MLR: Heteroskedasticity and related Stata options (Ch. 8, HD8), class 18
10. More specification and data issues in MLR: Functional Form Misspecification, Proxy Variables, Model with Random Slopes, Measurement Error (Ch. 9, HD9) classes 19–20
11. Simple Panel Data Estimation and related Stata options (Ch. 13, HD10), classes 21–22
12. Advanced Panel Data Estimation and related Stata options (Ch. 14, HD11), classes 22–24
13. Instrumental Variable Estimation, related Stata options, and final exam preparation (Ch. 15, HD12), classes 24–26

Material from appendices A, B, C, D, E, and G will also be used throughout the course.

Feedback I highly encourage students to provide me with feedback on how to further improve this course. I also conduct an informal and anonymous survey during the semester to seek feedback.

Computer policy I allow the use of laptop computers and other electronic devices in class, but only for course-related purposes such as making notes or reading electronic handouts. I do not allow the use of computers and other devices with Internet access during the exam.

Honor Code The College Honor Code will be observed. In particular, it is not acceptable to copy homework answers from other groups or from any other sources. Term papers should be original, cite all sources of data and ideas, and provide a do-file in the appendix as additional evidence of your original coding.

Disability Service William & Mary accommodates students with disabilities in accordance with federal laws and university policy. Any student who feels he/she may need an accommodation based on the impact of a learning, psychiatric, physical, or chronic health diagnosis should contact Student Accessibility Services staff at 757-221-2509 or at sas@wm.edu to determine if accommodations are warranted and to obtain an official letter of accommodation. For more information, please seewww.wm.edu/sas .

Required and Suggested Reading

Notation:
*– required reading
No asterisk – suggested reading for those who wish to learn more than is required to pass the course with an A. Also, suggested reading may help you with your project.

*All lecture slides, all writing on the graphic pen tablet/whiteboard, and all Stata codes discussed in class (slides and codes should be available on Blackboard)

*Wooldridge, J.M.. Introductory Econometrics. A Modern Approach. Sixth Edition. Cengage Learning. 2016.

Angrist, J. and Pischke J-S. 2009. Mostly Harmless Econometrics: An Empiricist's Companion.

Cameron, A. C. and Trivedi, P. Microeconometrics. Methods and Applications. Cambridge University Press, 2005.

Cameron, A. C. and Trivedi, P. Microeconometrics using Stata. Revised Edition. Stata Press, 2009.

Greene, W. Econometric Analysis. Pearson; 7 edition, February 13, 2011.

Little, R. J. A and Rubin, D. B., 2002. Statistical Analysis with Missing Data 2nd Edition. Wiley-Interscience; 2d edition.

Maddala, G.S., 1986. Limited-Dependent and Qualitative Variables in Econometrics. Econometric Society Monographs. Cambridge University Press; Revised ed. Edition, June.

Peter A. Savelyev, Benjamin Ward, Robert Krueger and Matt McGue. Health Endowments, Schooling Allocation in the Family, and Longevity: Evidence from US Twins. *Journal of Health Economics*, 81 (2022) 102554. Working paper almost identical to the published version: HCEO.

Wasserman, L. All of Statistics. Springer. 2003.

I reserve the right to alter the form and content of the course in order to adjust to the needs and level of students enrolled in the class.

Group Formation Questionnaire
(I need one form filled from each group.
Plus, I need one from each student, who has no group.)

1. Your name:

2. Have you formed a group of 2–4 students (including you) enrolled in this class?
   (Yes/No)   [If no, please skip to question 5]

3. Please list names of your group members (excluding you):

   (1)

   (2)

   (3)

4. Is your group open to admitting more members among those who found no group
   (to be assigned to your team by the professor)? (Yes /No) [skip to question 6]

5. Would you like to be assigned to a group by the professor? Yes / No

6. Your comments or suggestions (if any):