Quantitative Methods for Public Policy

Fall 2021

Instructors:	Room:	E-mail:	Office Hours:
Prof. Mihály Fazekas	C408	FazekasM@ceu.edu	Upon appointment
Teaching Assistant:			
Viktoriia Poltoratskaya		Poltoratskaya_Vikt@phd.ceu.edu	Upon appointment

Course Description

This course will introduce students to common quantitative research design and methods used for public policy analysis. Students will learn how to formulate practical research questions, find and construct appropriate datasets used for policy analysis, and use a variety of statistical tools to provide insight to important policy issues. Students will also learn how to use statistical software to manage and analyze various types of data. No prior knowledge of statistics is needed.

Learning Outcomes

By the end of the course students should be proficient in:
- ✓ Identifying and generating answerable research questions with direct policy implications;
- ✓ Using Stata to construct datasets and produce basic descriptive statistics;
- ✓ Using appropriate statistical analyses to answer specific research questions;
- ✓ Critically appraising quantitative research findings, and;
- ✓ Explain the limitations of observational data for making causal claims;
- ✓ Write clean, reusable, and reliable STATA code;
- ✓ Presenting analyses in a professional manner that is accessible to policy-makers.

Assessment

Grades will have two components:
1. Homework sets (50%): Homework will focus on practicing quantitative techniques learned in class. There will be 5 homework sets each worth 10 %. Due dates and topics are in the detailed syllabus below. Each homework set will be distributed a week before the deadline.
2. Bi-weekly quizzes (50%): Quizzes will be short, about 15 mins, multiple-choice tests of key concepts learnt in the previous classes. There will be 5 quizzes each worth 10%, dates are in the detailed syllabus below.

NOTE: Due to the large size of the class, there will be NO extensions granted for any of the homeworks. If you are ill and can provide medical documentation, weightings for your grade will be adjusted to reflect the missed assignment. All the assignments are strictly individual, no group work

or copying is permitted.

Course Readings

Mandatory

1. *Data Analysis for Business, Economics, and Policy* (2021) by Gábor Békés and Gábor Kézdi. Cambridge University Press.
2. *Essentials of Statistics for Business and Economics* (2015) by Anderson, Sweeney, Williams, Camm and Cochran (ASWCC).
3. *A Handbook of Statistical Analyses using Stata* (2004) by Sophia Rabe-Hesketh & Brian Everitt. Third Edition. CRC Press, London.

Note: Students are welcome to purchase earlier editions of the above book but must note that the material may differ slightly from what is presented in class.

Optional

Students, particularly those without a strong background in statistics or econometrics, are strongly encouraged to read the following:

1. *Naked Statistics: Stripping the Dread from the Data* (2014) by Charles Wheelan.
   An introduction to statistics with lots of intuition and examples. An easy-to-understand book that helps the reader learn how research questions are formulated and answered.
4. *Statistics Unplugged, 4th edition* (2013) by Sally Caldwell.
   A great companion to ASWCC for students who prefer more intuition.
5. *Introductory Econometrics: A modern approach, 6th edition* (2015) by Jeffery M. Wooldridge.
   An excellent resource for those who want more detail on regression analysis. One of the books that an empirical researcher or policy-maker should have on hand at all times. Please have a look at Appendix B and C if you are interested in a bit more detailed math behind the course material.
6. *Mastering 'Metrics: The Path from Cause to Effect* (2014) by Joshua Angrist and Jörn-Steffen Pischke.
   Angrist and Pischke are infamous for their econometrics books that put intuition and causal analysis front and center, using a host of real-world examples.
7. *Quantitative Social Science Methods* (2020) by Gary King.
   Collection of accessible and thorough videos from one of the best methods scholars in the field. Offering further examples, explanations and introduction to some more advanced topics.

Grading

CEU uses a system of letter grades and grade points for evaluating student work (please refer to the grade outline in the *CEU Student Records Manual*). Students who fail to submit work, or whose work fails to meet the minimum requirements for the assignment, will receive a grade of 'F'. **The lowest passing grade**

is C+. At the end of each course, course instructors distribute a detailed breakdown of the course grade components. ***Failing a mandatory (core) course results in termination from the program.***

Organization of the course

Essential and useful course resources and lecture slides will be available through the University electronic learning platform **Moodle.** These resources include required readings where necessary, homework, articles, platform to submit homework and guidance. Students should check the course site on Moodle on a regular basis as some required readings and homework will be uploaded as well as other important information. Students are encouraged to form groups in which they can discuss the readings before class and work on the homework.

Please note, that lecture slides are only crude outlines of the lectures, and they do not substitute notes taken in class or reading the textbook.

All comments and feedback concerning lectures, slides, selected readings etc. are welcome.

Schedule

Please note that this schedule is approximate and is subject to change. Importantly, we would like to offer 1-2 Stata labs for those who need additional support in acquiring the necessary coding skills. Dates to be agreed based on demand.

Some advice: Your success in this course will depend on keeping up with the material as it is presented. We strongly urge you not to fall behind because the material in the course is intensely cumulative. You will also benefit much more from the lectures if you read the assigned material before the class sessions.

No.	Date	Instr.	Topic
1	20-21 Sept	MF	**Introduction** Introduction to the course (why quants methods, aims, structure, etc); basics of research design such as theories, variables and operationalization, hypotheses; quantitative vs. qualitative research. Introduction to Stata. <u>Readings</u>: ASWCC Ch. 1 Rogoff-Reinhart scandal: summary: http://www.bbc.co.uk/news/magazine-22223190
2	27-28 Sept	MF	**Data Collection, Measurement, and Theory** Data collection methods including survey methods, sample vs. population; research designs and theory; measurement error. Advanced introduction to Stata. <u>Readings</u>: Adcock, R. and Collier D. (2001) "Measurement Validity: Toward a Shared Framework for Qualitative and Quantitative Research." American Political Science Review, 95(3), 529-546. Sophia Rabe-Hesketh & Brian Everitt (2004) A Handbook of Statistical Analyses using Stata. Third Edition. CRC Press, London. Chapter 1.
3	4-5 Oct	MF	**Data Preparation and Data types** Tidy data, source of variation, variable types, missing data, data cleaning Homework 1 (measurement and theory), due 8 October 5pm <u>Readings</u>: Békés-Kézdi (2021) Ch.2
4	11-12 Oct	MF	**Probability Distributions, Sampling and Hypothesis Testing** Basic probability, normal distribution, sampling distributions Null vs. alternate hypotheses, type I vs. type II error, one vs. two tailed tests. Homework 2 (data preparation and data types), due 15 October 5pm <u>Readings</u>: ASWCC Ch. 7-9
5	18-19 Oct	MF	**Exploratory Data Analysis and Descriptive Statistics** Measures of central tendency; measures of dispersion; describing distributions, widely used visualizations; check data quality, give context to subsequent analysis, and provide preliminary answers to simple questions. Quiz 1 (probability and sampling) <u>Readings</u>: Békés-Kézdi (2021) Ch.3
6	25-26 Oct	TA	**Comparison with Means and ANOVA** Comparisons of group means, one and two sample t-tests with equal and unequal variance; analysis of variance (ANOVA). Quiz 2 (descriptive stats) <u>Readings</u>: ASWCC Ch.10 Dávid-Barrett, Elizabeth & Fazekas, Mihály (2020) Anti-corruption in aid-funded procurement: Is corruption reduced or merely displaced? World Development. 132 (especially table 6 on simple unmatched group comparisons)
7	1-2 Nov	MF	**Introduction to Regression and Correlation** Correlation; simple regression model; model assumptions; significance testing. Implementation in Stata & replicating a research paper. Homework 3 (T-test, ANOVA), due 5 November 5pm <u>Readings</u>: ASWCC Ch. 12 Nicholas Charron, Carl Dahlström, Mihály Fazekas, and Victor Lapuente, (2017), Careers, Connections and Corruption Risks In Europe. Journal of Politics, 79(1). 89-104. Data&do files: https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/CSHNIS

8	8-9 Nov	MF	**Multiple Regression Basics** Multiple regression mode; interpreting coefficients; significance testing; multicollinearity; heteroscedasticity. Implementation in Stata & replicating a research paper. Quiz 3 (correlation) <u>Readings</u>: ASWCC Ch. 13 Nicholas Charron, Carl Dahlström, Mihály Fazekas, and Victor Lapuente, (2017), Careers, Connections and Corruption Risks In Europe. Journal of Politics, 79(1). 89-104. Data&do files: https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/CSHNIS
9	15-16 Nov	MF	**Further Issues with Multiple Regression** Functional form: using logs, quadratics, interaction terms; goodness of fit: r-squared, adjusted r-squared; predictions. Implementation in Stata & replicating a research paper. Homework 4 (Regression & Correlation) due 19 November 5pm <u>Readings</u>: Békés-Kézdi (2021) Ch.8 Nicholas Charron, Carl Dahlström, Mihály Fazekas, and Victor Lapuente, (2017), Careers, Connections and Corruption Risks In Europe. Journal of Politics, 79(1). 89-104. Data&do files: https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/CSHNIS
10	22-23 Nov	MF	**Inference with Multiple Regression** Using proxy variables; measurement error: dependent and independent variables, omitted variables, missing data, non-random samples, outliers. Implementation in Stata. Quiz 4 (regression I.) <u>Readings</u>: Békés-Kézdi (2021) Ch.8
11	29-30 Nov	MF	**Visual arguments** Principles of good data visualization, data visualization practice using Tableau Homework 5 (Multiple regression) due 3 December 5pm <u>Readings</u>: Edward Tufte (2001) The Visual Display of Quantitative Information. 2nd edition, Graphics Press. Chapter 2. Tableau introductory video (1. Tableau Public Overview): https://public.tableau.com/en-us/s/resources *Bring your favorite graph for discussion in class!*
12	6-7 Dec	MF	**Composite indicators** Composite score conceptualisation, imputation of missing data, weights, Principal Component Analysis. Implementation in Stata. Taster of what else lies on in this domain Quiz 5 (regression II.&visualisations) <u>Readings</u>: OECD (2008) Handbook on Constructing Composite Indicators, OECD, Paris, ch. 1. Sophia Rabe-Hesketh & Brian Everitt (2004) A Handbook of Statistical Analyses using Stata. Third Edition. CRC Press, London. Chapter 14.