**Research Design and Methods for Public Policy I**

**Fall 2020**

|  |  |  |  |
| --- | --- | --- | --- |
| Instructors: | Room: | E-mail: | Office Hours: |
| Prof. Mihály Fazekas | C408 | [FazekasM@ceu.edu](mailto:FazekasM@ceu.edu) | Upon appointment |
| Teaching Assistant: |  |  |  |
| T.B.A. |  |  |  |

**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.

**PLEASE NOTE: Due to the large size of the class, there will be NO extensions granted for any of the homework sets. If you are ill and can provide medical documentation, the weighting for your grade will be adjusted to reflect the missed assignment.**

**Course Readings**

**Mandatory**

1. *Essentials of Statistics for Business and Economics* (2015) by Anderson, Sweeney, Williams, Camm and Cochran (ASWCC).
2. *Data Analysis for Business, Economics, and Policy* (2021) by Gábor Békés and Gábor Kézdi. Cambridge University Press.
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.

1. *Statistics Unplugged, 4th edition* (2013) by Sally Caldwell.

A great companion to ASWCC for students who prefer more intuition.

1. *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.

1. *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.

1. *Quantitative Social Science Methods* (2020) by Gary King.

Collection of accessible and thorough videos from one of the best methods scholar 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.

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 | 28-30 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.  Readings:  ASWCC Ch. 1  Rogoff-Reinhart scandal: summary: <http://www.bbc.co.uk/news/magazine-22223190>  Sophia Rabe-Hesketh & Brian Everitt (2004) A Handbook of Statistical Analyses using Stata. Third Edition. CRC Press, London. Chapter 1. |
| 2 | 5-7 Oct | MF | **Data Collection, Measurement and Descriptive Statistics**  Types of variables; surveying, sample vs. population; measurement error, measures of central tendency; measures of dispersion; describing distributions, widely used visualizations. Introduction to Stata.  Readings:  ASWCC Ch. 2 & 3  Sophia Rabe-Hesketh & Brian Everitt (2004) A Handbook of Statistical Analyses using Stata. Third Edition. CRC Press, London. Chapter 2. |
| 3 | 12-14 Oct | MF | **Introduction to Probability Distributions**  Basic probability, normal distribution, sampling distributions.  Homework 1 (descriptive stats), due 16 October 5pm  Quiz 1  Readings:  ASWCC Ch.4, Ch. 6-8 |
| 4 | 19-21 Oct | MF | **Hypothesis Testing**  Null vs. alternate hypotheses, type I vs. type II tests, one vs. two tailed tests.  Homework 2 (probability and hypothesis testing), due 23 October 5pm  Readings:  ASWCC Ch.9 |
| 5 | 26-28 Oct | MF | **Exploratory data analysis**  Check data quality, give context to subsequent analysis, and provide preliminary answers to simple questions.  Békés-Kézdi (2021) Ch.3  Quiz 2 |
| 6 | 2-4 Nov | 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 3  Readings:  ASWCC Ch.10 |
| 7 | 9-11 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 13 November 5pm  Readings:  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 | 16-18 Nov | MF | **Multiple Regression Basics**  Multiple regression mode; interpreting coefficients; significance testing; multicollinearity; heteroscedasticity. Implementation in Stata & replicating a research paper.  Readings:  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 | 23-25 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 27 November 5pm  Readings:  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 | 30 Nov  -2 Dec | 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  Readings:  Békés-Kézdi (2021) Ch.8 |
| 11 | 7-9 Dec | MF | **Visual arguments**  Principles of good data visualization, data visualization practice using Tableau  Homework 5 (Multiple regression) due 11 December 5pm  Readings:  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 | 14-16 Dec | MF | **Composite indicators, reflections on the world of quantitative analysis/statistics/data science**  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  Readings:  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. |