Course Title: Data Management and Analysis for Public Policy

Instructor: Caitlin Brown

ECTS: 2

Term: Spring

Module: SFI

Relation to other courses:
Prerequisite: Research Methods I

Background and Overall Aim of the Course:

Managing and analyzing data is a key component of successful public policy. Data can be used to inform public policy on a range of levels, from providing important background information to evaluating the impact of a particular policy.

This course aims to provide students with the practical skills needed to adeptly construct, manage, and analyze data from a variety of sources. The course will use Stata as the primary programming language, though students may also follow along using another program of their choice (for example, R). The focus will be predominately on micro-level data, such as individual- or household-level data. The course requires little prior knowledge of statistical programming and is geared towards students from a wide range of backgrounds.

Learning Outcomes:

By the end of the course students should be proficient in:

- Finding and downloading data from a range of sources
- Merging different datasets together
- Data management skills such as cleaning and updating data
- Generate descriptive statistics and graphs describing the data
- Export results from statistical analysis to create professional tables

Learning Activities and Teaching Methods:

This course will be an applied practical class where students will learn the methods in-class as we progress throughout the course. The course will be part-lecture, where the basics are presented, and part "hands-on" work, where students put the methods learnt into practice.

Assessment:

Grades will be assessed based on the following:

Participation/Attendance:	30 %
Take-Home Exam:	70%

Course Content and Readings:

This course will be drawing from an array of sources. The key materials will be posted on Moodle for students to access. Students will also have access to Stata through one of CEU's computer labs.

For those looking for additional resources, the following are recommended:

- Alan C. Acock, A Gentle Introduction to Stata, 5th edition, (Stata Press: 2016).
- *A Handbook of Statistical Analyses using Stata* (2004) by Sophia Rabe-Hesketh & Brian Everitt. Third Edition. CRC Press, London.
- *Mastering 'Metrics: The Path from Cause to Effect* (2014) by Joshua Angrist and Jörn-Steffen Pischke.

The course schedule is as follows:

Topic 1: Data for public policy

- How data can be used to make public policy decisions.
- Types of data, where to locate data, how to download and access data.
- Introduction to Stata: commands, do-files and results

Topic 2: Variable and dataset management

- Data management in Stata: converting between string and numeric data types, handling of dates, importing and exporting data, creating new variables with gen and egen, use of variable and value labels, recoding variables

Topic 3: Creating new datasets for policy analysis

- How to merge different datasets in Stata, reshaping data

Topic 4: Descriptive analysis and basic statistical inference

- Descriptive statistics, hypothesis tests involving means and proportions, cross-tabulation, basic multiple regression

Topic 5: Visually displaying data in Stata

- Using the graph twoway command: combining graphs, specifying labels, axes, ticks etc.
- Distribution plots: histogram, qnorm, quantile

Topic 6: Creating tables for publication using Stata

- Exporting results from Stata to tables: esttab and outreg commands