

**More Advanced Econometric Programming using Stata**Eric de Bodt – [eric.debodt@nhh.edu](mailto:eric.debodt@nhh.edu)

- **Prerequisites**

This lecture is a follow-up from 2021 “An Introduction to Econometric Programming using Stata” lecture. I assume therefore that students are used to Stata programming constructs, data manipulation and standard statistical analysis commands. I can’t emphasize this enough because if you are new to Stata, most of the coding work that we will do will appear to you as black magic !

For students wishing a memory refresh, I provide the videos corresponding the following topics:

- Lecture 1 : Anatomy of a research project in Stata
- Lecture 2 : The Building Blocks
- Lecture 3 : More Building Blocks
- Lecture 4 : Data management
- Lecture 5 : More on programming

The detailed content of each video is in Appendix 1 and the videos are available at the OneDrive location:

<https://1drv.ms/u/s!AmE8oP2EftLkkqYC2K3wyQJnqzwF3A?e=IM57Un>

- **Lecture outline**

Lecture 1 : Panel data estimators

Lecture 2 – Diff-in-diff estimators –theory

Lecture 3 – Diff-in-diff estimators for event staggered through time – Stata code

Lecture 4 – Randomized inference

Lecture 5 : From regression to Lasso and Elastic Nets – a first step towards machine learning

Bonuses:

- Difference-in-differences using Stata `didregress` and `xididregress`
- Binned scatter plots
- Fixed Effects models - Theory

- **Lecture organization**

Like in 2021, there will be one class by week. The Monday, I will release to you a video and an assignment. The Friday, a Zoom session will be organized to allow you either to raise questions (when the topic is more theoretical) or to study together my solution to the assignment (when it comes to Stata code). The final assessment will be based on the collection of responses provided to the five assignments, submitted through the NHH Wiseflow system.



## Appendix 1

### 2021 “An Introduction to Econometric Programming using Stata” Support Videos

#### Lecture 1 : Anatomy of a research project in Stata

- Why writing Stata code ?
- Stata coding guidelines
- Optimization:
  - o Compiled commands
  - o Implicit loops
- Structuring Stata code for research projects
  - o Data management
  - o Descriptive statistics
  - o Models estimation
  - o Tables/figures generation
- The first seven : `cls / clear / pause on / set more off / set linesize / cd / version`
- Structuring the code: `program / args`
- My best friends: `display / pause / log / timer / trace`

#### Lecture 2 : The Building Blocks

- directories : `cd / pwd`
- data types : `byte / int / long / float / double / str / strL`
- missing data : `missing() / mvencode + mvdecode`
- handling errors : `capture / r()`
- controlling the display : `quietly`
- local / global macro (literal, expression, nested, evaluation, ++)
- macro extended functions
- scalars / matrices
- namespaces
- command classes: `r_class / e_class / c_class / s_class`
  - o stored results
  - o return / ereturn
- `preserve / restore`
- implicit loops: `by & statsby prefix, _n & _N constants`
- explicit loops: `foreach / for values`
  - o Loops on macros, variables, etc
  - o Loops on observations (`x[i]`)

### Lecture 3 : More Building Blocks

- varlist (wildcards \* and -) / numlist
- creating & updating variables: generate / replace / egen / egenmore
- selecting observations
  - o keep / drop
  - o if / in / cond() / inlist() / inrange()
- data conversion functions and commands:
  - o real() / string() / recode() / autocode() / egen cut() / egen group()
  - o tostring / destring / recode / encode / decode / recode
- factor variables / leads and lags (calendar dates) versus x[\_n-1]
- managing dates
  - o Date data types
  - o Date formats : %td / %tc / ...
  - o From date literal to date value : td() / tw() / ...
  - o Date conversion functions: date() / mdy () / ...
- memory management: compress / memory / task manager

### Lecture 4 : Data management

- importing data: infile / insheet / import excel
- augmenting data: append / merge
- restructuring data:
  - o reshape
  - o fillin / cross / joinby
  - o xpose / stack / separate
- testing: assert / duplicates / isid / unique / codebook

### Lecture 5 : More on programming

- Collections and Tables (new Stata 17)
- Figures:
  - o Graph types
  - o Essential twoway options
- Stata matrix : svmat / mkmat / matrix list & operators
- postfile
- frame (new Stata 16)