Effective Programming Practices for Economists

Version Control and collaboration with Git and Github

Introduction and Overview

Janoś Gabler and Hans-Martin von Gaudecker

What is git

- Definition: A distributed version control system
- Practical: Something that lets you ...
  - ... go back in time
  - ... keep track of different versions of your code
  - ... try out changes without fear of breaking anything
  - ... document your project in a commit history

What is GitHub

- Definition: GitHub is a platform and cloud-based service for software development and version control using Git
- Practical: Something that lets you
  - ... collaborate efficiently with coauthors
  - ... review changes to your code before accepting them
  - ... manage coding and research projects via issues and pull requests
  - ... build a portfolio for a highly payed tech job

Why researchers need git and GitHub

- Most papers are written by groups
- Empirical and computational projects become more complex
- The publication process entails multiple rounds of revisions
- Reproducibility and transparency are key

The learning journey

To quote John von Neumann:

> Young man, in ~~mathematics~~ **git** you don't understand things. You just get used to them.

- Learning git is a long journey
- Payoffs are enormous and once you learned it you will not want to go back!
- The best way to learn it is:
  - Learn git on your local computer first
  - Only use the shell!
  - Practice, practice, practice!
- Don't be afraid of an ugly and convoluted commit history, we won't deduct anything for that

Notes about the video series

- We will use an existing video series to learn git
- Examples are from web development
  - Works the same for any other plain text file
- A few things have changed since they were recorded
  - `master` has been renamed to `main`
  - The Atom editor has been replaced by VS Code
  - On Windows, the Powershell has seen improvements, so forget about commander