Repository Layout

May 24, 2025

This PDF is generated from `writeup/tex/README.tex`. It documents the folder hierarchy and how each directory feeds into the others.

Contents

1 Directory Structure

`data/` Raw inputs from external sources live in `data/raw`. Scripts under `src/` write processed panels to `data/processed`. Small samples kept under version control reside in `data/samples`.

`src/` Stata build scripts that turn raw files into analysis-ready datasets. Every script sources `src/globals.do` for consistent paths and writes its output to `data/processed`.

`spec/` Self-contained empirical specifications. Each `.do` file loads prepared panels from `data/processed`, runs a model, and exports both tidy tables and raw diagnostics into `results/`.

`py/` Small Python utilities that post-process Stata results. They merge standard errors, generate figures, and tidy output before it reaches the paper.

`results/` Generated artefacts from all estimation scripts. The folder is split into `raw/`, `cleaned/`, and `figures/`. Clean tables are later copied into `writeup/`.

`writeup/` Contains the LaTeX source of the paper along with a Makefile. The build rules collect the cleaned tables from `results/cleaned` and compile the final PDF.

2 Interdependencies

The directories form a linear pipeline:

1) **Raw data** enters under `data/raw`.
2) Stata scripts in `src/` transform these files into structured panels stored in `data/processed`.
3) Empirical models in `spec/` read the processed panels and write their outputs to `results/`.
4) Python helpers in `py/` further clean these outputs so that the `writeup/` build can include them directly.
5) The LaTeX sources within `writeup/` assemble the final report using the cleaned tables and generated figures.

Each stage depends only on the outputs of the previous one. Rebuilding a panel or specification updates downstream folders without affecting unrelated components.