Replication package for "Efficiency and Redistribution in Environmental Policy: An Equilibrium Analysis of Agricultural Supply Chains"

This document describes the contents of the replication package and the steps required to produce the tables and figures in the main text and online appendix.

1. Contents

The replication package consists of three main folders: `code`, `data`, `output`. Each of these folders contains the following subfolders:

- `code/`
  - `0_dataprep`: contains R scripts that process the raw data files and produce the clean datasets used in the analysis.
  - `1_descriptives`: contains R scripts that produce the tables and figures used in the descriptive analysis of the paper (i.e., the stylized facts from the main text and the descriptives from the online appendix).
  - `2_estimation`: contains R scripts that run the econometric analysis, produce all estimation-related tables and figures, and set up the parametrization of the quantitative model.
  - `3_simulations`: contains Julia scripts that solve the model equilibrium under different conduct assumptions and policy environments. These scripts also produce all simulation-related tables and figures (model fit, policy counterfactuals, welfare analysis).
- `data/`
  - `sources.xlsx`: an Excel file containing descriptions of all raw data files and their precise source. Each sheet in this workbook corresponds to a data category (e.g., prices, land use, production).
  - The data folder contains several subfolders organized by data category (e.g., `data/production`). Each of these subfolders has a `raw` subfolder and a `clean` subfolder:
    - `raw`: contains data files exactly as they are downloaded from their source (described in `sources.xlsx`).
    - `clean`: contains processed versions of the data (produced by the scripts in `code/0_dataprep`).
- `output/`
  - `figures`: contains all figures from the main text.

- `figures/appendix`: contains all figures from the online appendix.
- `tables`: contains all tables from the main text.
- `tables/appendix`: contains all tables from the online appendix.

2. Steps for replication

The naming of the subfolders within `code` and the individual scripts they contain indicate the order in which they must be executed to successfully replicate all results.

To replicate the tables and the figures in the main text and the online appendix you will need to execute the following four codes in this order:

- `code/0_dataprep/0_start.R`
- `code/1_descriptives/0_start.R`
- `code/2_estimation/0_start.R`
- `code/3_simulations/0_start.jl`

The `start` scripts in each subfolder will automatically set the directory to the location of the script and will run all codes in the subfolder sequentially.

Recommendations for successful replication:

- To avoid problems related to incompatible directories in R, it is highly recommended to use RStudio to run the R scripts. The `0_start.R` scripts are written in such a way that when they are executed in RStudio the directory will be automatically set to the script's location.
- To avoid problems with missing R packages, it is highly recommended to install the pacman package before running any scripts. This package will automatically install any required packages when running a specific R script.
- Once you have installed the pacman package, all the `0_start.R` scripts will run seamlessly on RStudio.