The Notebooks were constructed to minimize complexity of a replication. At the beginning of the Jupyter Notebooks you just can insert the path to the folders on your local machine. Necessary libraries (e.g., Numpy or Pandas) might need to be installed if you run the code on your local machine.

1. Construction of the Independent Variables

Constructing the Independent Variables

Files needed:

- From Pretransformed Data folder
- From Map Data folder

Files created:

- Independent Variables.csv

2. Construction of the Dependent Variable

Merge assigned Occupational Titles (HISCO) and Skill Levels (HISCLASS) with Occupational Data and calculate the expected value of each skill level per county. Then save it on the local machine.

Note: There is a second notebook to create the dependent variable for a different assignment of skills to occupations for a robustness check.

Files needed:

- For Dependent Variable of Main Regression and most Robustness Checks:
    - Classification Data folder
    - Control Variables from Pretransformed Data folder
    - Wage data from the ifo Data folder

Files created:

- Main_Dep.csv (assigned skill per county – grouped)
- Total_Var.csv (assigned skill per profession per county – not grouped)
- Other dependent variables for robustness checks

3. Main Regression

Files needed:

- Dependent Variable from the Saving the Dependent Variables folder
- Independent Variables from the Saving the Independent Variables folder

- Steam Engine Data from the Pretransformed Data folder

4. Robustness Checks

Files needed:

- Dependent Variable from the Saving the Dependent Variables folder
- Independent Variables from the Saving the Independent Variables folder
- Steam Engine Data from the Pretransformed Data folder

5. Exclusion Restriction

Files needed:

- Dependent Variable from the Saving the Dependent Variables folder
- Independent Variables from the Saving the Independent Variables folder
- Data from the ifo Data folder

6. Maps

Files needed:

- Dependent Variable from the Saving the Dependent Variables folder
- Data from the Classification Data folder
- Data from the Map Data folder

7. EDA

Files needed:

- Dependent Variable from the Saving the Dependent Variables folder
- Data from the Pretransformed Data folder
- Data from the Classification Data folder
- Data from the Saving the Independent Variables folder

Files created:

- Overview Professions.csv
- Overview Variables.csv

8. HISCO Machine

Function that gives you the skill levels of HISCO codes. Can be used to compare if assigned HISCO codes have the same skill level.