Problem Set 4

ECON833, Prof. Jason DeBacker
Gyumin Kim

Q1. Write a function to take the "raw" data and cleans it (i.e., it should accept and return a DataFrame object where the input is the raw and the output has the correct sample selection and necessary variables).

A1 : Answer for Question 1 is written in ProblemSet4 kim.py file.

Q2. Write 2 unit tests for your data cleaning function.

A2 : Answer for Question2 is written in Unit test.py file. I attached a screenshot in the last part of the PDF file because I wrote down my codes for question 2 and 5 in the same .pu file and ran pytest -v.

Q3. Estimate the following model via a Maximum Likelihood Estimator separately for t = 1971, 1980, 1990, 2000:

A3 :

$$\ell(\boldsymbol{\beta}, \sigma^2) = -\frac{n}{2}\ln(2\pi\sigma^2) - \frac{1}{2\sigma^2}(\mathbf{y} - X\boldsymbol{\beta})^\top(\mathbf{y} - X\boldsymbol{\beta}). \tag{1}$$

This is the equation for the likelihood function that I estimate.

Q4. Interpret the coefficient β_1. How do the returns to education change over time in these data?

A4 : In 1971, the coefficient on years of education was 0.0665, indicating that each additional year of schooling raised wages by about 6.7 percent. In 1980, the return to education was nearly identical at 0.0660, or 6.6 percent, showing little change from the early 1970s. By 1990, however, the return to education had increased substantially to 0.0955, implying a 9.6 percent wage premium for each additional year of schooling. This upward trend continued into 2000, where the coefficient rose further to 0.1103, suggesting that returns to education reached approximately 11 percent per year of schooling. Overall, the results show that the economic payoff to education was relatively stable in the 1970s and 1980s but rose sharply during the 1990s and 2000s, highlighting the growing importance of schooling in determining wages.

Q5. Write a unit test that checks that your MLE is correct. You can do this by comparing your MLE to the results from an OLS regression.
A5 :

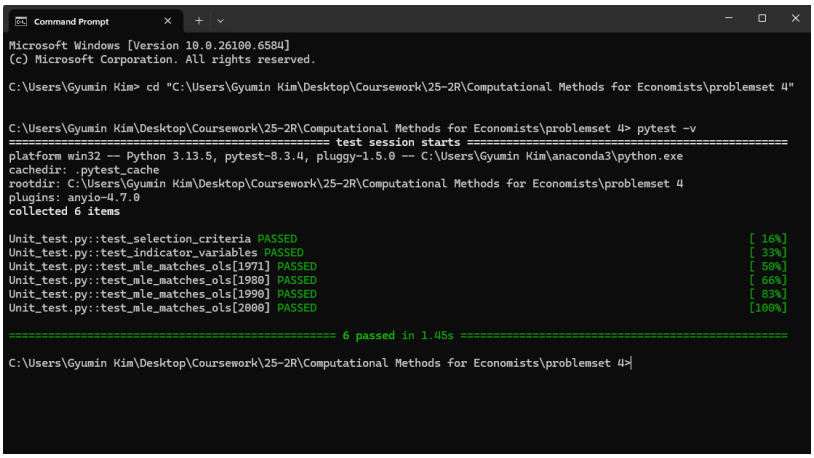

Figure 1: Results for Unit test required in question 2 and 5