

Econ 8186: Homework #1

Due back: Tuesday, January 31st by Noon

Instructions: You are allowed to work in groups of at most 2 students and submit together. Please make sure to specify the members of the group clearly. Please submit a clear report explaining your findings that is self contained along with any code (Stata, R, Fortran) and dataset (when applicable). All files should be in one zip file and the code should run without changing directory names or other modifications. Please make every effort to produce a complete package so that I can understand what you did, how you did it, and what you found.

1. Consider a standard consumption-savings problem with stochastic income and a risk free bond (what is also called an “income-fluctuations problem”):

$$\begin{aligned} V(a_t; y_t) &= \max_{c_t} (u(c_t) + \delta \mathbb{E}[V(a_{t+1}; y_{t+1}) | y_t]) \\ \text{s.t.} \quad c_t + a_{t+1} &= y_t + (1+r)a_t, \\ \log y_t &= \rho \log y_{t-1} + \eta_t, \\ a_t &\geq \underline{a}. \end{aligned} \tag{1}$$

$$a_t \geq \underline{a}. \tag{2}$$

Derive the Euler equation for each of the following preference specifications. For (a), derive the Euler equation first by using the FOC and envelope condition and a second time by using the variational approach we discussed in class. You can use either approach for the remaining cases.

- (a) $U = \frac{(c_t^i - \theta c_{t-1}^i)^{1-\alpha}}{1-\alpha}$ where c_t^i denotes individual i 's consumption
- (b) $U = \frac{(c_t^i - \theta \bar{C}_{t-1})^{1-\alpha}}{1-\alpha}$ where \bar{C} is aggregate consumption
- (c) $U = \frac{(c_t^i - \theta \bar{C}_t)^{1-\alpha}}{1-\alpha}$

2. This question is intended to familiarize you with the US Current Population Survey data, which is the most commonly used household survey dataset in the United States. You will be asked to download certain variables, clean the data, apply sample selection criteria, and produce some statistics, tabulations, and report some regression results. We will use data from the 2012 March Current Population Survey, ASEC supplement (so income is for year 2011). Download these data from IPUMS, which is the most convenient source. You may want to compare your results with your classmates before submitting to catch mistakes.

- Download the following variables for each individual: age, education (last degree completed), gender, and state of residence. Also download the income variables INCWAGE (last year's total wage and salary income), “usual weekly hours”, and “weeks worked”.
- Create a sample by applying the following sample selection criteria:
 - Keep if age is between 25 and 60, inclusive.
 - Drop individuals with missing data on incwage, usual weekly hours, weeks worked, age, education, and gender. Keep if state is missing.
 - Drop if INCWAGE is less than \$2,000 (2012 dollars).

- Drop if hourly wage > \$500, annual hours < 50 hrs. Hourly wage is INCWAGE divided by the product of “usual weekly hours” and “weeks worked”.
- (a) Tabulate summary statistics for the raw data you downloaded and for the final sample you created with the selection criteria just described: report the sample size; and the mean, median, standard deviation, min, max, skewness, kurtosis, and interquartile range for all the variables you downloaded. (Prepare two tables, one for each sample, with the same format.)
- (b) Run the following version of the Mincer wage regression, where $y_i \equiv \log(Y_i)$ is log income (incwage):

$$y_i = d_{\text{age}} + d_{\text{educ}} + d_{\text{gen}} + d_{\text{state}}, \quad (3)$$

where d_j indicates a dummy for variable j , and the regression is run in a single cross section (2012). Use one dummy for each age (omit dummy for age 25 for normalization). Education is the last degree completed and is a categorical variable with 4 possible values: (i) high school degree; (ii) some college; (iii) college degree; and (iv) graduate education, or at least some post-baccalaureate education. (The fifth dummy—for less-than-highschool education—is omitted for normalization.)

- Report:

- the seniority gap (between age 25 and 55),
- the gender gap
- the college premium
- the name of the states with lowest and highest average log wage and the wage level in each (based on state dummy).

- (c) Now run an expanded Mincer wage regression adding interaction terms:

$$\begin{aligned} \log w_{it} = & a_0 d_{\text{age}} + a_1 d_{\text{educ}} + a_2 d_{\text{gen}} + a_3 d_{\text{age}} \times d_{\text{educ}} \\ & + a_4 d_{\text{gen}} \times d_{\text{educ}} + a_4 d_{\text{gen}} \times d_{\text{age}} + a_5 d_{\text{state}} + \text{error} \end{aligned}$$

- Report:

- the seniority gap (between age 25 and 55) for men and women by education level.
- the average gender gap (how do you calculate?), and the gender gap for college graduates at ages 30, 40, and 50.
- the college premium, separately for each gender, at age 45.
- the name of the states with lowest and highest average log wage and the wage level in each (based on state dummy).

- (d) Discuss how your results in parts b and c compare.