Problem Set #2b

Danny Edgel
Econ 715: Econometric Methods
Fall 2021

December 10, 2021

The attached file, functions.jl, includes all functions used in this problem set, including an OLS function. edgel_ps2.tex includes the code the executes the commands for the problem set. Using these files, the coefficient for education is derived as 0.151 .

To obtain a conditional average treatment effect (CATE) for increasing education from 12 years to 16 years, I first subset the data to only include the observations with education equal to either 12 or 16 years. Then, I generated binary variable $T_i = \mathbb{1}\{edu_i = 16\}$ and ran OLS on the following specification:

$$Y_i = \beta_0 + \beta_1 T_i + \beta_2 T_i X_i + \beta_3 X_i^2 + \beta_4 X_i + \beta_5 X_i^2 \varepsilon_i$$

Where X_i is years of experience. Then, the CATE for each year of experience is given by:
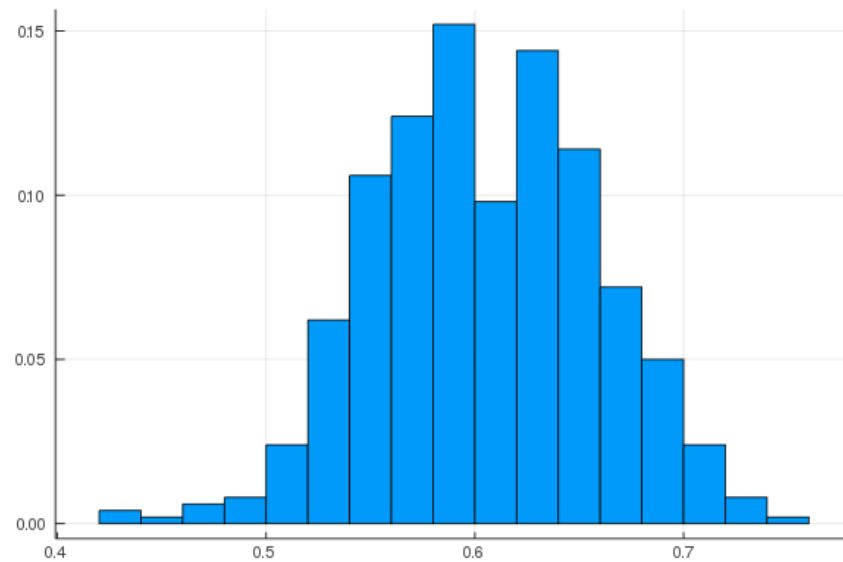
$$\tau(X) = \beta_1 + \beta_2 X + \beta_4 X^2$$

Which is plotted below, along with a 95% confidence interval.
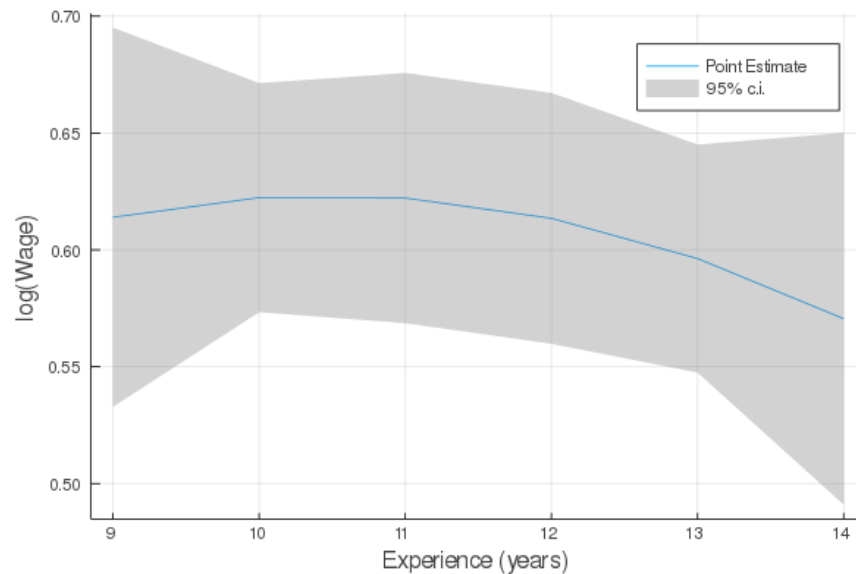
Using the sample shares of years of experience, the average treatment effect (ATE) is estimated as 0.606 .

Using a single sample of 400 observations, the naive ATE is estimated as 0.451 . A histogram of ATE estimates from 500 samples of 400 observations is shown below.[1]


Repeating this process, but calculating the CATE and averaging over all years of experience, the ATE is estimated as 0.448 , and the new histogram is:

---

[1] All samples were taken without replacement.
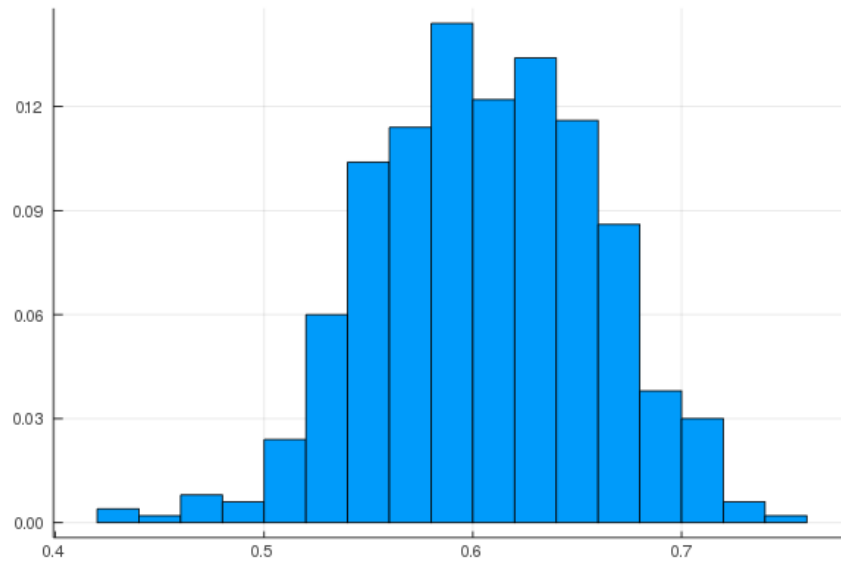
The estimate from question 1 (from OLS) falls far below the lower tail of each set of estimates. The CATE estimate from question 2b, however, falls in or near the modal bin of each histogram.

Using a propensity score as requested, the ATE estimate is 0.458 , and the histogram for 500 simulations is below. As you can see, the estimates vary much more than they do using the other two methods. However, the OLS estimate is still well below the minimum estimate, and estimate from 2(b) is above the median (and the modal bin).
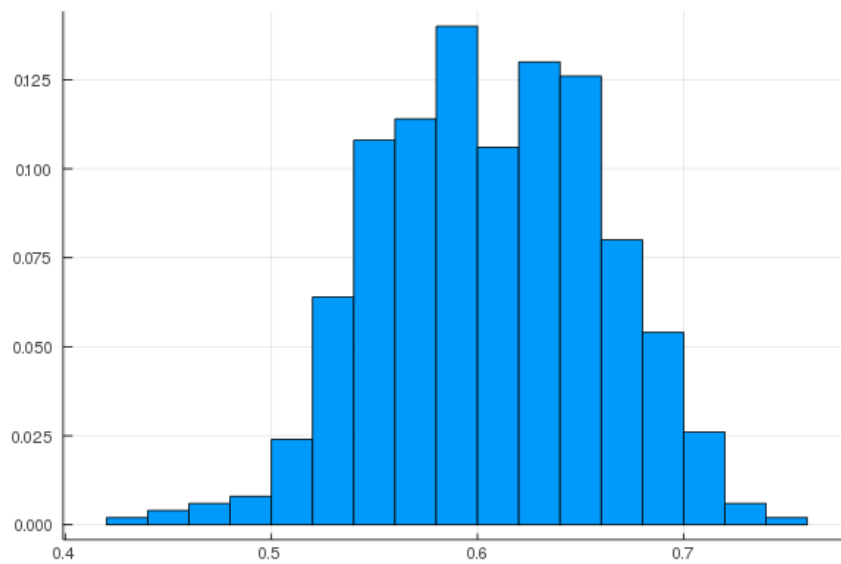
The three methods above yield different results with similar means. I would prefer the first method, which estimates the CATE for each experience level, then averages across experience. I prefer this method because it has the lowest variance and is the most theoretically robust, since the first figure (from question 1) shows that the treatment effect does, indeed, vary with experience.

When we draw samples of 400 for each of the problems above, we should use the same samples for each problem (through obviously not for each simulation) to control for variation in estimates that is due exclusively to sampling, which is not particularly interesting variation for comparing estimation methods.