20210 Summer 2025: Problem Set 4
Econometrics ♡
Instructor(s): Murilo Ramos

Agustín Esteva ♡
aesteva@uchicago.edu
Due Date: July 29, 2025

Problem 1

Stock and Watson, Exercises *6.11, *7.4, *7.8.(a)-(b)

Problem 2

Consider the following model of the determinants of wages:

$$\log(\text{wage}) = \beta_0 + \beta_1 \text{educ} + \beta_2 \text{pareduc} \times \text{educ} + \beta_3 \text{experience} + \beta_4 \text{experience}^2 + U,$$

where:

- **experience** = years of individual work experience
- **pareduc** = sum of the mother's and father's years of education
- **educ** = years of own education

(a) What is the percent return on wages from another year of own education? Does it depend on the level of own education? Does it depend on parents' education? Does it depend on the level of work experience?

> SOLUTION: Solving,
> $$\frac{\partial \log \text{wage}}{\partial \text{educ}} = \beta_1 + \beta_2 \text{paraeduc}$$
> We know from the SLR that
> $$\%\Delta Y = 100(\beta_1 + \beta_2 \text{paraeduc})\Delta \text{education}$$
> and so it depends on parents education only! ♡

(b) What is the percent return on wages from another year of work experience? Does it depend on the level of own education? Does it depend on parents' education? Does it depend on the level of work experience?

> SOLUTION: One can take partials to discover that
> $$\%\Delta Y = 100(\beta_3 + 2\beta_4 \text{experience})\Delta \text{experience}$$
> so it only deepends on experience! ♡

(c) Using 722 observations, the following OLS estimates were obtained:

$$\hat{\beta}_0 = 5.65, \quad \hat{\beta}_1 = 0.047, \quad \hat{\beta}_2 = 0.00078, \quad \hat{\beta}_3 = 0.22, \quad \hat{\beta}_4 = -0.005$$

    i. Compute the estimated return to an additional year of own education if both parents have high school education (pareduc = 24), and if both parents have college degrees (pareduc = 32). Use these results to interpret the estimated interaction between own education and parents' education.

SOLUTION: From part (a), we plug in the numbers and find the percent return to be

- (24) An increase of 6.574%

- (32) An increase of 7.196%

Thus, more parent education is associated with a higher increase in the return of wages given a higher number of your own education. ♡

ii. Compute the estimated return to an additional year of experience if the individual has zero years of experience (experience $= 0$), and if the individual has 20 years of experience (experience $= 20$). Use these results to interpret $\hat{\beta}_4$.

SOLUTION: From part (b), we plug in

- (0) An increase of 22%

- (20) An increase of 2%

Thus, β_4 is the sensitivity of the change in wages depending on the change experience from current experience. In this case, this β_4 is lowkey not good, as it points to diminishing returns. ♡

iii. How would the OLS estimates of β_3 and β_4 change if experience was measured in months instead of years?

SOLUTION: Just divide by 12 bruv:

$$\beta_0' = \beta_0 \quad \beta_1' = \beta_1 \quad \beta_2' = \beta_2 \quad \beta_3' = \frac{\beta_3}{12} \quad \beta_4' = \frac{\beta_4}{144}$$

♡

Problem 3

Levitt and Venkatesh investigate Chicago street gangs by modelling the determinants of the wages paid to "foot soldiers" in the gangs. They estimated the following regression:

$$\text{wage} = \beta_0 + \beta_1 \text{war} + \beta_2 \text{large} + U,$$

where:

- **wage** = the hourly wage paid to the foot soldiers

- **war** =
$$\begin{cases} 1 & \text{if the gang is currently involved in a gang war} \\ 0 & \text{otherwise} \end{cases}$$

- **large** =
$$\begin{cases} 1 & \text{if the gang is "large"} \\ 0 & \text{otherwise} \end{cases}$$

and found that $\hat{\beta}_0 = 1.83$, $\hat{\beta}_1 = 1.3$, $\hat{\beta}_2 = 4.07$.

(a) According to their estimates, what is the average wage paid to a foot soldier in a small gang that is not at war?

> SOLUTION: Plug in war $= 0$ and large $= 0$ to see that
>
> $$\hat{\text{wage}} = \hat{\beta}_0 = 1.83$$
>
> ♡

(b) Why did they include a dummy variable for the gang being large but not also for the gang being small?

> SOLUTION: Just plug in 0 or 1 depending on size of gang. ♡

(c) How would you modify their regression to allow for the effects of a gang war on wages to be different for members of large versus small gangs? What is the reference group in your suggested model?

> SOLUTION: Consider
>
> $$\text{wage} = \beta_0' + \beta_1' \text{war} + \beta_2' \text{large} + \beta_3' \text{large} \cdot \text{war} + U'$$
>
> reference group is still small gang not at war. ♡

(d) Explain the interpretation of each of the coefficients in your model.

SOLUTION: β'_0 same as before, β'_1 is the difference in wages between a small gang at war and at peace, β'_2 is the difference in wages between a warring gang that is big and one that is small, β'_3 is the difference of the differences coefficient- which measures how much the wages change depending on size and war ♡

Problem 4

Prove the missing steps of the FWL theorem from our lecture:

(a) $\text{cov}(X_j^*, X_l) = 0$ for all $l \neq j$

> SOLUTION: Consider that if
>
> $$X_j = \text{BLP}(X_j \mid X_{-j}) + U_{X_j},$$
>
> then we have showed in class that since the BLP minimizes mean square error, then
>
> $$\mathbb{E}[U_{X_j}] = 0,$$
>
> but
>
> $$U_{X_j} = X_j - \text{BLP}(X_j \mid X_{-j}) = X_j^*,$$
>
> and so we are done since
>
> $$\begin{aligned}
\text{Cov}(X_j^*, X_\ell) &= \mathbb{E}[X_j^* X_\ell] - \mathbb{E}[X_j^*]\mathbb{E}[X_\ell] \\
&= \mathbb{E}[\mathbb{E}[X_j^* X_\ell \mid X_\ell]] \\
&= \mathbb{E}[X_\ell \mathbb{E}[X_j^* \mid X_\ell]] \\
&= 0
\end{aligned}$$
>
> ♡

(b) $\text{cov}(X_j^*, X_j) = \text{Var}(X_j^*)$

> SOLUTION: We compute
>
> $$\begin{aligned}
\text{Cov}(X_j^*, X_j) &= \text{Cov}(X_j^*, X_j^* + \text{BLP}(X_j \mid X_{-j})) \\
&= \text{Var}(X_j^*) + \text{Cov}(X_j^*, \alpha_0 + \alpha_1 X_1 + \cdots + \alpha_{j-1} X_{j-1} + \alpha_{j+1} X_{j+1} + \cdots + \alpha_k X_k)
\end{aligned}$$
>
> But we know by part (a) that the second term vanishes. ♡

(c) $\text{Cov}(X_j^*, U) = 0$

> SOLUTION: Recall that a basic assumption of the model is that $\text{Cov}(X_j, U) = 0$ for any $j \in [k]$. Thus,
>
> $$\begin{aligned}
\text{Cov}(X_j^*, U) &= \text{Cov}(X_j - \sum_{i \neq j} \alpha_i X_i, U) \\
&= \text{Cov}(X_j, U) - \sum_{i \neq j} \alpha_i^2 \text{Cov}(X_i, U)
\end{aligned}$$

$$= 0$$

by the assumption. ♡

where $X_j^* = X_j - \mathrm{BLP}(X_j|X_{-j})$ and $u = Y - \mathrm{BLP}(Y|X)$.

Problem 5

(* You will need to download the dataset for this problem from the course webpage.) The dataset has 4 variables in the following order: `college` (an indicator for completing college), `hs` (an indicator for completing high school, but not college), `wage` (average hourly wage), and `fem` (an indicator for female).

Consider the following model of the determinants of income:

$$\text{wage} = \beta_0 + \beta_1 \text{fem} + \beta_2 \text{college} + \beta_3 \text{hs} + U$$

(a) Show that the regressors are not perfectly colinear in this dataset. (Hint: regress one regressor on the other regressors).

(b) Assume the model is causal. What is the interpretation of U? Is it uncorrelated with the regressors? Interpret each of the coefficients in the model.

(c) Assume the model above is the BLP(wages|fem, college, hs). What is the interpretation of U? Is it uncorrelated with the regressors? Interpret each of the coefficients in the model.

(d) Estimate a model that allows the partial correlation (coefficients of the BLP) between college and wages to depend on gender, and the partial correlation between high school completion and wages to depend on gender. Report your table of estimates.

(e) Test at the 5% significance level whether the partial correlation between college and wages depends on gender. Explain your results.

(f) Test at the 5% significance level whether the partial correlation between high school or college and wages depends on gender. Explain your results.