

Homework 2: OLS and Probit

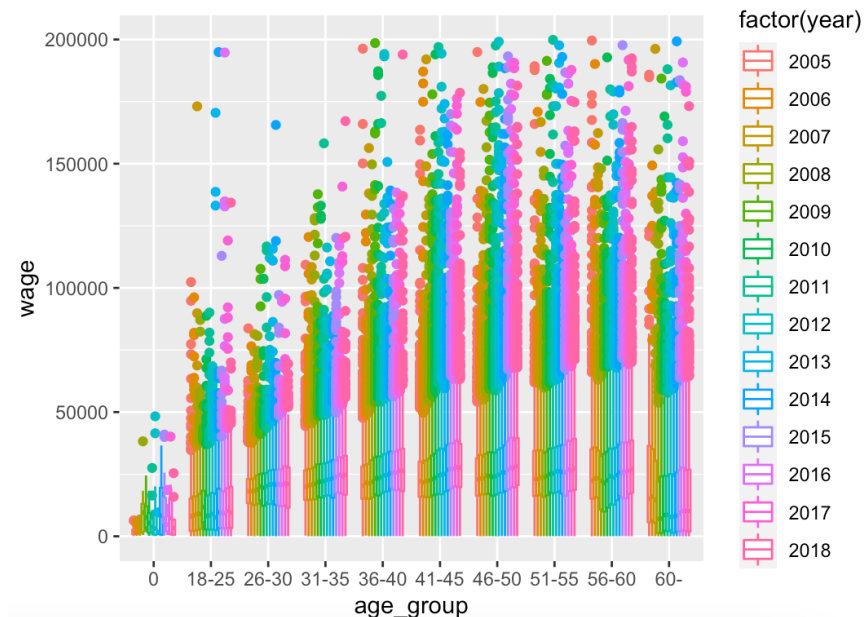
Runling Wu

1 Exercise 1 OLS estimate

- The correlation between Y and X is -0.1788512.
- The intercept is 12568.7505 and slope is 276.5593.
- The standard errors of β
 - Using the standard formulas of the OLS, we have se for slope as 16.79083.
 - Using bootstrap with 49 replications, we have the se for slope as 15.75011.
 - Using bootstrap with 499 replications, we have the se for slope as 15.73966.
 - The difference between standard formulas of OLS and bootstrap is that bootstrap allows for existence of outliers while keep estimation consistent.

2 Exercise 2 Detrend Data

- I plot the wage distribution the wage of each age group across years



- Use boxplot to plot the full distribution. IF we look at the distribution, the lowest point seems to be increasing over 2005 to 2018. Thus, there is an increasing time trend of wage distribution across all age groups. It makes perfect sense since we are only plotting nominal wages across the past few years, without accounting for the inflation (overall change in the price index level).
- After including a time fixed effect, the estimated coefficients will increase. Since

	OLS FE	OLS
age	326.31*** (5.02)	276.56*** (16.79)
(Intercept)		12568.75*** (717.18)
R ²	0.03	0.03
Adj. R ²	−0.05	0.03
Num. obs.	156858	10495

*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$

3 Exercise 3 Numerical Optimization

- The likelihood of the probit of being employed is 6582.155339.
- Optimize the model and interpret the coefficients. On the lefthand side, the result is from GLM package. On the righthand-side, the result is from manual optimization. β does not have structural meaning, we only accept the sign of age, which is negative. It means that as age increases, we will expect to see the probability of being employed decreases.

	R: GLM : est	R: own : est
(Intercept)	3.8291873	3.82928037
age	−0.0678642	−0.06786624

- We cannot, since probit model only include determinants that affect whether people choose to work or not, like wage shifters, but we cannot directly add wage.

4 Exercise 4 Discrete choice

- Interpret and compare the estimated coefficients. How significant are they? They are all significant at 1% level. We only accept the sign of age, which is negative. It means that as age increases, we will expect to see the probability of being employed decreases. For all the year dummies, we find that their coefficients are positive, thus we could conclude that years have positive effect on labor participation.

	Logit	Probit	PLM
age	−0.12*** (0.00)	−0.06*** (0.00)	−0.02*** (0.00)
year_2005	7.03*** (0.04)	3.57*** (0.02)	1.54*** (0.00)
year_2006	7.05*** (0.04)	3.58*** (0.02)	1.55*** (0.00)
year_2007	7.09*** (0.04)	3.61*** (0.02)	1.55*** (0.00)
year_2008	7.11*** (0.04)	3.61*** (0.02)	1.55*** (0.00)
year_2009	7.04*** (0.04)	3.56*** (0.02)	1.54*** (0.00)
year_2010	7.04*** (0.04)	3.57*** (0.02)	1.54*** (0.00)
year_2011	7.10*** (0.04)	3.60*** (0.02)	1.55*** (0.00)
year_2012	7.09*** (0.04)	3.58*** (0.02)	1.55*** (0.00)
year_2013	7.04*** (0.04)	3.56*** (0.02)	1.54*** (0.00)
year_2014	7.10*** (0.04)	3.59*** (0.02)	1.55*** (0.00)
year_2015	7.09*** (0.04)	3.57*** (0.02)	1.55*** (0.00)
AIC	153824.70	159808.12	
BIC	153946.57	159930.00	
Log Likelihood	−76900.35	−79892.06	
Deviance	153800.70	159784.12	
Num. obs.	190296	190296	190296
R ²			0.78
Adj. R ²			0.78

*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$

5 Exercise 5 Marginal Effects

- the marginal effect of the previous probit and logit models

	Probit	Logit
age	−0.0239798931	−0.0270037760
year_2005	−0.0005987002	−0.0123718775
year_2006	0.0027009278	−0.0077177930
year_2007	0.0027009278	0.0013091766
year_2008	0.0153576583	0.0053754071
year_2009	−0.0037254472	−0.0112710272
year_2010	−0.0016308407	−0.0093489434
year_2011	0.0085657438	0.0028369538
year_2012	0.0035060435	−0.0002095949
year_2013	−0.0070816538	−0.0095949843
year_2014	0.0042998617	0.0024206837

- The standard errors of the marginal effects

	Probit	Logit
age	0.0000960	0.0001175
year_2005	0.0062930	0.0067747
year_2006	0.0062156	0.0066614
year_2007	0.0060856	0.0064922
year_2008	0.0060863	0.0064675
year_2009	0.0061417	0.0065689
year_2010	0.0060491	0.0064528
year_2011	0.0059587	0.0063054
year_2012	0.0058891	0.0062308
year_2013	0.0060386	0.0064188
year_2014	0.0059518	0.0062711