

Homework 2: OLS and Probit
Julian Sauvage
Feb 4th, 2022

Exercise 1

- Correlation between Age and Wage:

```
> square_root %*% covar_matrix %*% square_root
      [,1]      [,2]
[1,]  1.0000000 -0.1788512
[2,] -0.1788512  1.0000000
```

- Regression Coefficients:

- o Intercept coefficient (β_0) = 12568.705
- o Age coefficient (β_1) = 276.559

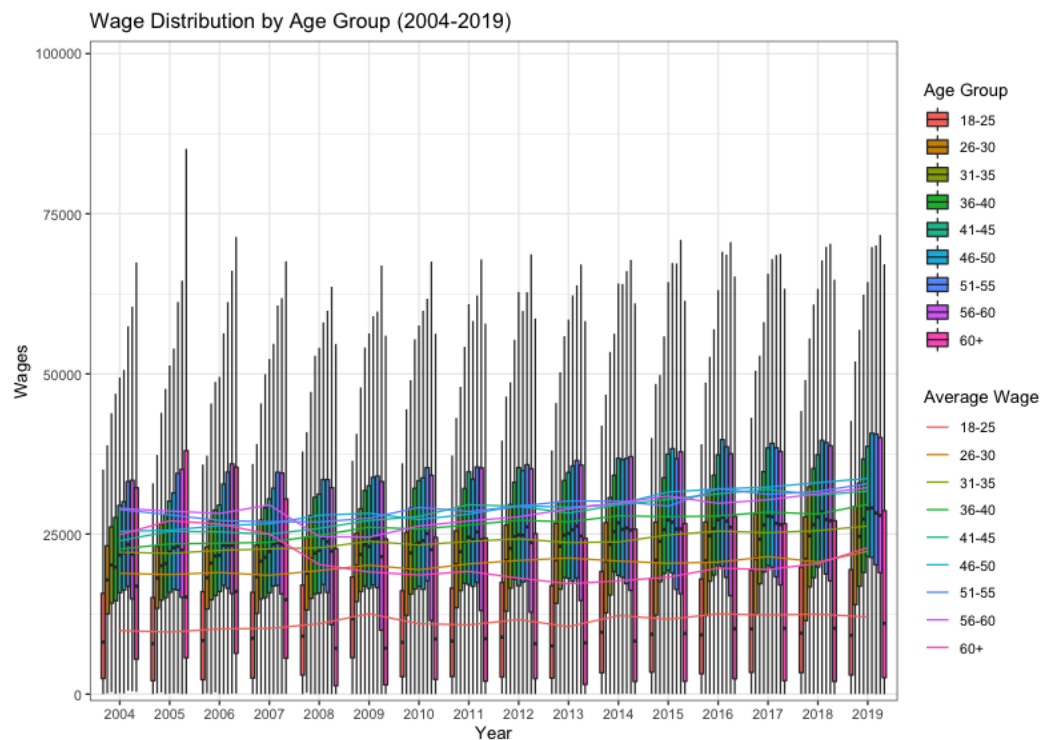
- Standard Errors:

```
Manual: est Manual: sd Bootstrap49: est Bootstrap49: se Bootstrap499: est Bootstrap499: se
1 12568.7505 717.49086 12644.1412 566.19408 12590.0883 583.55418
age 276.5593 16.79803 274.3578 15.75011 276.0287 15.73966
> |
```

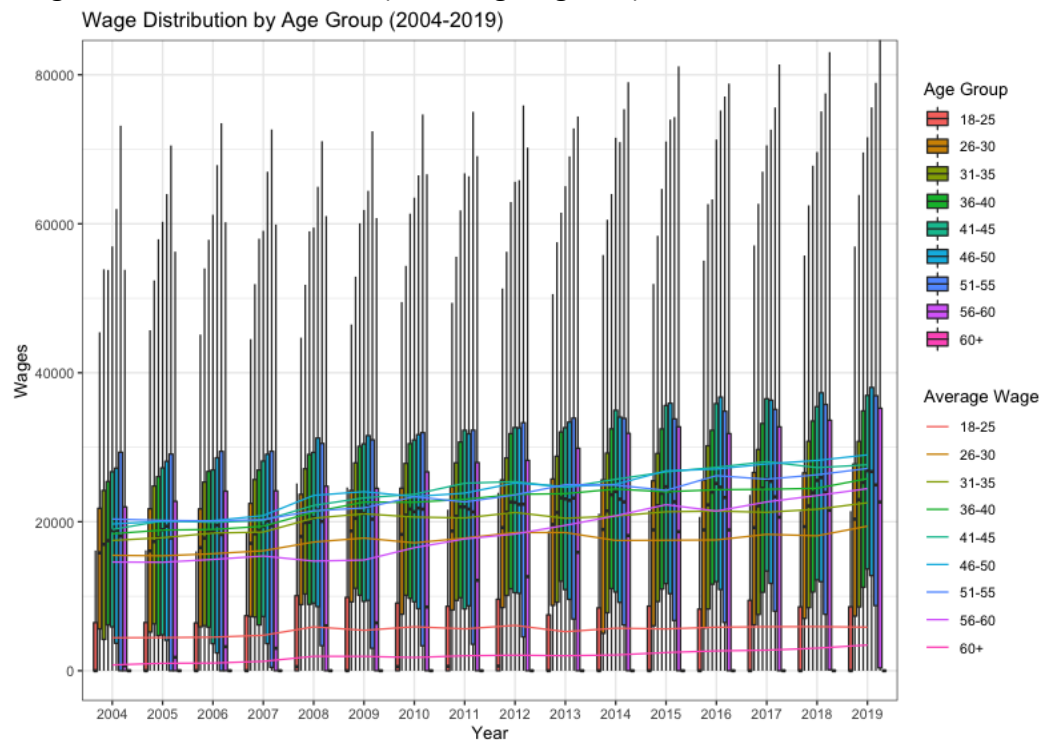
Using the standard OLS formula is more efficient than using bootstrap to calculate the standard errors. As the number of replications increases, the bootstrap estimates become more accurate, however it is clear this convergence occurs slowly.

Exercise 2

Wage of each age group across years (excluding wage = 0):



Wage Distribution over time (including wage = 0):



Excluding observations where wage is 0, produces some interesting effects. Most noticeable is the much smaller “whiskers”, which are a result of the lower variation at the bottom of the

distribution. More interesting however is the effect on the wage distribution of 60+ age group. Their average wages more than double, closely approximating the wages of the “26-30” age group. We also see an interesting trend around 2007-2008, where the average wages of the “56-60” and “60+” groups dip significantly. This is likely attributable to the “Great Recession” and may be because the higher paid individuals in these groups left the workforce. Alternatively, these groups may have simply moved to lower-paying jobs/industries.

Time fixed effect:

- After including a time fixed-effect, the estimated coefficient for age is negative, and all coefficients become smaller in magnitude.

Exercise 3

Interpreting Coefficients:

```

-      Built-in Probit: est Built-in Probit :se Own Probit : est Own Probit :se
(Intercept)      1.045238785      0.060229466      1.042389218      0.058446571
age               0.006900094      0.001451705      0.006964405      0.001404556

```

The coefficient on age (~0.0069) is positive, suggesting that an increase in age also increases the predicted probability that an individual is employed. The intercept suggests that the predicted probability that one is employed with an age = 0 is 85% ($F(1.04) = 0.85$), which is illogical. This is likely a result of the simple regression utilized, and the fact that retired individuals were excluded from this analysis. If retired individuals were included, we might want to include an age² term to represent the quadratic relationship between age and employment status.

Labor market participation:

We cannot estimate the same model while including wages as a determinant because wages are determined after an individual decides to participate in the labor market and becomes employed. Moreover, wages are only available to those who are employed, and thus we would not see any variation of wages amongst those who are unemployed.

Exercise 4

```

> est1
Built-in Probit: est Built-in Probit :se Own Probit : est Own Probit :se
(Intercept)      0.77343624      0.0172471182      0.77365878      0.0168008789
age               0.01212493      0.0004166431      0.01212438      0.0004044834

> est2
Built-in Logit: est Built-in Logit :se Own Logit : est Own Logit :se
(Intercept)      1.16227647      0.0326889274      1.15926018      0.0326747551
age               0.02492306      0.0008086262      0.02500079      0.0008082767

> est3
Built-in Linear: est Built-in Linear :se Own Linear : est Own Linear :se
(Intercept)      0.801344376      3.176802e-03      0.826297200      8.103233e-04
age               0.002305359      7.404533e-05      0.001698165      8.618996e-06

```

Coefficients:

The coefficients are all significant but there are interesting differences across the 3 models and compared to the previously estimated probit model from exercise 3. In the probit

we estimated here, the intercept is below 1, and the effect of age is almost double. This suggests that age has a stronger positive effect on employment than previously noted.