Econ 613 Homework 2
Hannah Marsho

Exercise 1: OLS Estimate

a) 0.14349
b) 230.9923
c) (i) 14.8774
(ii) 49 replications:

	coefficient: estimate	coefficient: std dev	bootstrap (49): estimate	bootstrap (49): std dev
(Intercept)	14853.9230	635.46022	14149.8497	596.08216
age	213.4264	14.63047	232.1823	15.72919

499 replications:

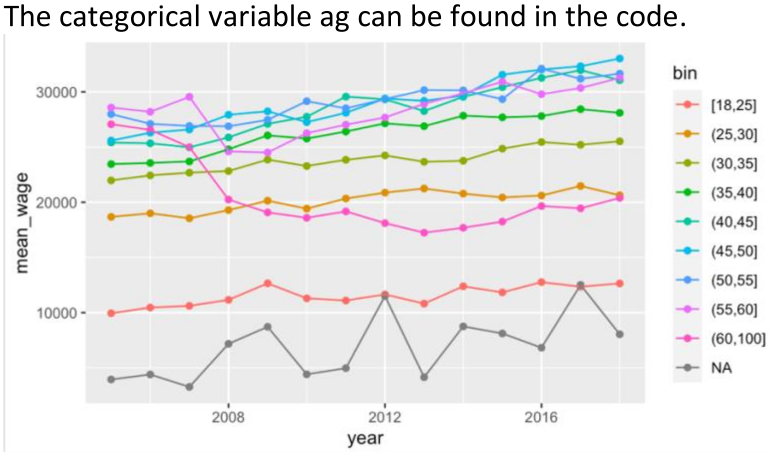
	coefficient: estimate	coefficient: std dev	bootstrap (499): estimate	bootstrap (499): std dev
(Intercept)	15212.4054	644.00898	14146.9951	622.14890
age	211.0909	14.85426	230.9198	16.38194

The OLS strategy applies the standard formulas to the actual data, finding the true values. The Bootstrap strategy creates multiple samples (as many samples as replications) that are representative of the data/population. It then calculates the coefficient and standard error estimates for each sample/replication and then creates a composite coefficient and standard error estimate based on that.

Both methods give similar results. As we see above, adding more bootstrap replications helps our coefficient and standard error estimates become better approximates of the model. In other words, the values get closer to the true values when we add replications.

Exercise 2: Detrend Data

a) The categorical variable ag can be found in the code.



b)
Older age groups tend to have higher mean wages across years. Over the years, the wages tend to either stagnant or rise for each age group.

```
Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept) -5.359e+05  2.782e+04  -19.26   <2e-16 ***
age          3.060e+02  4.319e+00   70.85   <2e-16 ***
year         2.722e+02  1.384e+01   19.67   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

c)
After including the time fixed effect, the age coefficient grew larger (and the standard error decreased).

Exercise 3: Numerical Optimization

a) The cleaned dataset can be found in the code.
b) The function can be found in the code. The likelihood is 2079.097.
c) The optimization can be found in the code.

```
      V1         V2       V3
1.092757 0.01307698 2079.097
```

Above are the optimized coefficients. The intercept is 1.092757, the age coefficient is 0.01307698, and the likelihood value associated with this specification is 2079.097. Probit coefficients have no intrinsic meaning. The only thing we can take away from them is what their sign indicates about the relationship between the associated variables. The positive age coefficient here means that being older is, on average all else equal, associated with a larger probability of being employed.

d) We cannot estimate the same model including wages as a determinant of labor market participation. This is because there is an underlying relationship between wages and labor market participation. If your wage is 0, then there’s a very strong, if not definite, chance that you are unemployed. There is also a two-way relationship between wages and labor market participation. Due to this, we cannot sufficiently outline a probability relationship between the two in this way.

Exercise 4: Discrete Choice

a) The cleaned dataset can be found in the code.
b) The functions and optimizations can be found in the code.
c) The functions and optimizations can be found in the code.

Probit Optimized Model:

```
      V1       V2        V3       V4       V5
8.809346 -9.08887 0.5621098 7.848381 110372.2
```

Above are the optimized coefficients. The intercept is 8.809346, the age coefficient is -9.08887, the year coefficient is 0.5621098, and the likelihood value associated with this specification is 110372.2.

Logit Optimized Model:

V1	V2	V3	V4	V5
8.809346	-9.08887	0.5621098	7.848381	110372.2

Above are the optimized coefficients. The intercept is 8.809346, the age coefficient is -9.08887, the year coefficient is 0.5621098, and the likelihood value associated with this specification is 110372.2.

Linear Probability Optimized Model:

V1	V2	V3	V4	V5
3.352805	-9.011688	3.902105	-2.734769	110372.2

Above are the optimized coefficients. The intercept is 3.352805, the age coefficient is -9.011688, the year coefficient is 3.902105, and the likelihood value associated with this specification is 110372.2.

Across the different models, the coefficients are similar. The age coefficient is negative, meaning that, as you get older, you are less likely to work. This is different from our result in Exercise 3. For each model, a zero-hessian matrix was returned, so I was unable to receive the variances to calculate the significances. The parts of the for loops relating to the hessian matrices are left in to show my ideas/progress, but they are "commented out" with the # sign.

Exercise 5: Marginal Effects

I wanted to write a nice, neat function that would return all the necessary values for Exercise 5 well. However, to do so, I had to use the glm() function in my solution, as I was unable to figure out how to write the function without using it to find the initial parameters.

a and b) Probit Marginal Effects and Standard Errors of Probit Marginal Effects:

	Marginal Effect	Standard Error of Marginal Effect
(Intercept)	6.432932385	4.150655e-01
age	0.002140708	7.979613e-05
year	-0.003145615	2.066545e-04

Logit Marginal Effects and Standard Errors of Logit Marginal Effects:

	Marginal Effect	Standard Error of Marginal Effect
(Intercept)	6.490828242	5.297117e-01
age	0.002243804	8.141724e-05
year	-0.003189967	2.633701e-04