[redacted]

Name: Ben Chu

[redacted]

ben-chu@berkeley.edu

**Problem 1. Multivariate linear regression**

**a)**

```
set obs 1000

gen u = rnormal(0,5)

gen x1 = rnormal(0,1)

gen x2 = exp(x1)

gen y = 2 + 4*x1 - 6*x2 + u
```

**b)**

```
reg y x1
reg y x2
```

Name: Julian Bertalli
SID: 3033589368
Section ID: 103
jjbertalli@berkeley.edu

Name: Ben Chu
SID: 3032040857
Section ID: 110
ben-chu@berkeley.edu

```
. reg y x1

      Source |       SS           df       MS      Number of obs   =     1,000
-------------+----------------------------------   F(1, 998)       =    305.12
       Model |  30102.3749         1  30102.3749   Prob > F        =    0.0000
    Residual |  98458.6631       998  98.6559751   R-squared       =    0.2341
-------------+----------------------------------   Adj R-squared   =    0.2334
       Total |  128561.038       999  128.689728   Root MSE        =    9.9326

------------------------------------------------------------------------------
           y |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          x1 |  -5.652941   .3236203   -17.47   0.000    -6.287995   -5.017887
       _cons |  -7.738369   .3142795   -24.62   0.000    -8.355093   -7.121644
------------------------------------------------------------------------------

. reg y x2

      Source |       SS           df       MS      Number of obs   =     1,000
-------------+----------------------------------   F(1, 998)       =   2889.71
       Model |  95558.6372         1  95558.6372   Prob > F        =    0.0000
    Residual |  33002.4008       998  33.0685379   R-squared       =    0.7433
-------------+----------------------------------   Adj R-squared   =    0.7430
       Total |  128561.038       999  128.689728   Root MSE        =    5.7505

------------------------------------------------------------------------------
           y |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          x2 |   -4.61281     .08581   -53.76   0.000    -4.781199   -4.444422
       _cons |  -.3314056   .2260605    -1.47   0.143     -.775014    .1122029
------------------------------------------------------------------------------
```

Name: Julian Bertalli
SID: 3033589368
Section ID: 103
jjbertalli@berkeley.edu

Name: Ben Chu
SID: 3032040857
Section ID: 110
ben-chu@berkeley.edu

Looking at random data generated in Stata, the true constant $\beta_0 = 2$, while the true coefficient for x1 = 4 and the true coefficient for x2 = -6. A sample of n>30 should be considered larger enough since the sample size is 1000. Despite the sample size being 1000, the OLSE does not converge to the actual parameters for either x1 or x2.

The differences arise between the population sloped and the estimated slopes for a few reasons. Since the first OLSE assumption does not hold, $E[u_i|x_i] = 0$, then the null hypothesis ($H_0$): $\beta_1 = 4$ would have to get rejected in the case of a two-side alternative in the first case even though 4 is the true value. The first case has an error term, $u_i$ of $-6x_{2i} + u_i$, but the conditional expectation of this term given $x_{1i}$ does not equal zero according to the definition of $x_{2i}$.

The same thing happens with the second case.

**c)**

reg y x1 x2

```
. reg y x1 x2
```

| Source | SS | df | MS |
|---|---|---|---|
| Model | 102790.205 | 2 | 51395.1024 |
| Residual | 25770.8333 | 997 | 25.8483784 |
| Total | 128561.038 | 999 | 128.689728 |

| | |
|---|---|
| Number of obs | = 1,000 |
| F(2, 997) | = 1988.33 |
| Prob > F | = 0.0000 |
| R-squared | = 0.7995 |
| Adj R-squared | = 0.7991 |
| Root MSE | = 5.0841 |

| y | Coef. | Std. Err. | t | P>\|t\| | [95% Conf. | Interval] |
|---|---|---|---|---|---|---|
| x1 | 4.152341 | .2482524 | 16.73 | 0.000 | 3.665184 | 4.639498 |
| x2 | -6.029263 | .1136973 | -53.03 | 0.000 | -6.252376 | -5.806149 |
| _cons | 2.02335 | .2444688 | 8.28 | 0.000 | 1.543618 | 2.503082 |

Name: Julian Bertalli
SID: 3033589368
Section ID: 103
jjbertalli@berkeley.edu

Name: Ben Chu
SID: 3032040857
Section ID: 110
ben-chu@berkeley.edu

The new regressors from the Stata output in c) are more accurate than in part b). The new numbers are extremely close to the true values. For example, the true value for x1 is 4, while the regression value is 4.152341. The true value of x2 is -6, while the regression value is -6.029263. The true value of the constant is 2, while the regression gives an output of 2.02335. The standard errors (SE) are smaller in the new model that only contains x1 model, going from .3236203 in the old to .2482524 in the new. Since the new regression contains an additional regressor, the $R^2$ in c) will be greater than the $R^2$ in d) since $R^2$ is nondecreasing. Adding a regressor increases $R^2$. The Stata output generated shows that the new regressor added in part c) is a relevant variable. This new variable is explanatory in relation to the data. For this reason, the adjusted $R^2$ will also increase after the new regression in part c).

**d)**

```
gen x3 = 1 + x1 - x2 + rnormal(0,0.5)
reg y x1 x2 x3
```

Name: Julian Bertalli
SID: 3033589368
Section ID: 103
jjbertalli@berkeley.edu

Name: Ben Chu
SID: 3032040857
Section ID: 110
ben-chu@berkeley.edu

```
. gen x3 = 1+ x1 - x2 + rnormal(0,0.5)

. reg y x1 x2 x3
```

| Source | SS | df | MS |
|---|---|---|---|
| Model | 102790.763 | 3 | 34263.5876 |
| Residual | 25770.2754 | 996 | 25.8737705 |
| Total | 128561.038 | 999 | 128.689728 |

| | | |
|---|---|---|
| Number of obs | = | 1,000 |
| F(3, 996) | = | 1324.26 |
| Prob > F | = | 0.0000 |
| R-squared | = | 0.7995 |
| Adj R-squared | = | 0.7989 |
| Root MSE | = | 5.0866 |

| y | Coef. | Std. Err. | t | P>\|t\| | [95% Conf. | Interval] |
|---|---|---|---|---|---|---|
| x1 | 4.200567 | .4117583 | 10.20 | 0.000 | 3.392554 | 5.008581 |
| x2 | -6.077313 | .3464244 | -17.54 | 0.000 | -6.757119 | -5.397508 |
| x3 | -.0476044 | .3241786 | -0.15 | 0.883 | -.6837559 | .5885471 |
| _cons | 2.071899 | .4112515 | 5.04 | 0.000 | 1.26488 | 2.878918 |

Do the changes in the OLS estimates, standard errors, the $R^2$, and the adjusted R make sense to me? By including x1, x2, and x3, there is imperfect multicollinearity between the regressors since the linear combination between x1 and x2 can explain the majority of the variations contained within x3.

Additionally, despite the OSLE assumptions holding true with the actual population parameter of $\beta_3 = 0$, the standard errors (SE) actually increase for each regressor due to multicollinearity. Since a third regressor was added for part d), the $R^2$ is the same as in part c) since $R^2$ is nondecreasing. The $R^2$ is .7995 in both cases, which makes sense since it did not decrease as a new regressor was added.

Conversely, the adjusted $R^2$ is not necessarily nondecreasing. In this case, the adjusted $R^2$ is actually decreasing since the new added regressor does not have an actual effect or impact on the dependent variable. The new regressor has a relatively small explanatory value on the dependent

Name: Julian Bertalli
SID: 3033589368
Section ID: 103
jjbertalli@berkeley.edu

Name: Ben Chu
SID: 3032040857
Section ID: 110
ben-chu@berkeley.edu

variable within the regression. The adjusted $R^2$ falls from .7991 in part c) to .7989 in part d). Even though this is only a very small decrease after the regressor was added, it does make sense to me since adjusted $R^2$ is capable of increasing or decreasing.

$R^2$ part c): .7995
$R^2$ part d): .7995

Adjusted $R^2$ part c): .7991
Adjusted $R^2$ part d): .7989

**e)**

```
set obs 1000

gen u = rnormal(0,5)

gen x1 = rnormal(0,1)

gen x2 = exp(x1)

gen y = 2 + 4*x1 - 6*x2 + u

regress y x1 x2

summ y x1 x2
```

Name: Julian Bertalli
SID: 3033589368
Section ID: 103
jjbertalli@berkeley.edu

Name: Ben Chu
SID: 3032040857
Section ID: 110
ben-chu@berkeley.edu

```
.
. regress y x1 x2
```

| Source | SS | df | MS |
|---|---|---|---|
| Model | 77357.4768 | 2 | 38678.7384 |
| Residual | 23354.6973 | 997 | 23.4249723 |
| Total | 100712.174 | 999 | 100.812987 |

| | | |
|---|---|---|
| Number of obs | = | 1,000 |
| F(2, 997) | = | 1651.18 |
| Prob > F | = | 0.0000 |
| R-squared | = | 0.7681 |
| Adj R-squared | = | 0.7676 |
| Root MSE | = | 4.8399 |

| y | Coef. | Std. Err. | t | P>\|t\| | [95% Conf. | Interval] |
|---|---|---|---|---|---|---|
| x1 | 4.198106 | .2625917 | 15.99 | 0.000 | 3.682811 | 4.713402 |
| x2 | -6.094774 | .132733 | -45.92 | 0.000 | -6.355243 | -5.834306 |
| _cons | 1.914831 | .26126 | 7.33 | 0.000 | 1.402148 | 2.427513 |

correlate y x1 x2

```
. correlate y x1 x2
(obs=1,000)
```

| | y | x1 | x2 |
|---|---|---|---|
| y | 1.0000 | | |
| x1 | -0.5270 | 1.0000 | |
| x2 | -0.8418 | 0.7998 | 1.0000 |

Name: Julian Bertalli
SID: 3033589368
Section ID: 103
jjbertalli@berkeley.edu

Name: Ben Chu
SID: 3032040857
Section ID: 110
ben-chu@berkeley.edu

**Problem 2. Teaching ratings**

**a)**

reg course_eval beauty, robust

```
. reg course_eval beauty, robust

Linear regression                               Number of obs     =        463
                                                F(1, 461)         =      16.94
                                                Prob > F          =     0.0000
                                                R-squared         =     0.0357
                                                Root MSE          =     .54545
```

| course_eval | Coef. | Robust Std. Err. | t | P>\|t\| | [95% Conf. | Interval] |
|---|---|---|---|---|---|---|
| beauty | .1330014 | .0323189 | 4.12 | 0.000 | .0694908 | .1965121 |
| _cons | 3.998272 | .0253493 | 157.73 | 0.000 | 3.948458 | 4.048087 |

The estimated slope is .1330014, which is the marginal effect that beauty has on the category course_eval score, which means that according to this sample, instructors who have a higher "beauty" rating get higher instructional rating score. With a p-value of 0.000, the value is statistically significant at the 1% significance level (as well as the 5% and 10% levels). Low p-value = high significance.

**b)**

reg course_eval beauty intro onecredit female minority nnenglish, robust

Name: Julian Bertalli
SID: 3033589368
Section ID: 103
jjbertalli@berkeley.edu

Name: Ben Chu
SID: 3032040857
Section ID: 110
ben-chu@berkeley.edu

```
. reg course_eval beauty intro onecredit female minority nnenglish, robust

Linear regression                               Number of obs     =        463
                                                F(6, 456)         =      17.03
                                                Prob > F          =     0.0000
                                                R-squared         =     0.1546
                                                Root MSE          =     .51351
```

| course_eval | Coef. | Robust Std. Err. | t | P>\|t\| | [95% Conf. | Interval] |
|---|---|---|---|---|---|---|
| beauty | .16561 | .0315686 | 5.25 | 0.000 | .1035721 | .2276478 |
| intro | .011325 | .0561741 | 0.20 | 0.840 | -.0990673 | .1217173 |
| onecredit | .6345271 | .1080864 | 5.87 | 0.000 | .4221178 | .8469364 |
| female | -.1734774 | .0494898 | -3.51 | 0.001 | -.2707337 | -.0762212 |
| minority | -.1666154 | .0674115 | -2.47 | 0.014 | -.2990912 | -.0341397 |
| nnenglish | -.2441613 | .0936345 | -2.61 | 0.009 | -.42817 | -.0601526 |
| _cons | 4.068289 | .0370092 | 109.93 | 0.000 | 3.995559 | 4.141019 |

When adding additional variables to control for the type of course and professor characteristics, the impact of beauty on course_eval score becomes higher: .1330014 in a) to .16561 in part b).

Since the 95% confidence interval of .1035721 to .2276478 contains the value from a): .1330014, we are unable to reject the notion that the coefficient is the same as it is in a). That is: we are unable to reject the null hypothesis that the effect of beauty is the same as it is in part a).

As a result, the results from the regression in a) may suffer from moderate omitted variable bias (based on the results from the regression).

Although there is not a large change in the beauty coefficient, .1330014 in a) to .16561 in part b), the $R^2$ does have a large increase from part a) to part b). The $R^2$ in part a) is .0357, while the $R^2$ in

Name: Julian Bertalli
SID: 3033589368
Section ID: 103
jjbertalli@berkeley.edu

Name: Ben Chu
SID: 3032040857
Section ID: 110
ben-chu@berkeley.edu

part b) is .1546. This does not prove causality because many variables may alter the fit of the $R^2$ within the context of this problem, such as enthusiasm, experience, and charisma.

On average, instructors teaching intro classes (mainly large Freshman and Sophomore courses) will receive slightly higher instructional ratings (.011325).

On average, instructors teaching one-credit classes will receive higher instructional ratings (.6345271).

On average, female instructors will receive lower instructional ratings in comparison to male instructors (-.1734774).

On average, minority instructors will receive lower instructional ratings in comparison to white instructors (-.1666154).

On average, non-native English speaking instructors will receive lower instructional ratings in comparison to native English speaking instructors (-.2441613).

Additionally, instructors who are perceived as "beautiful" will receive higher scores by an average of about .16516 points according to the regression.

The constant/baseline score for an instructor is 4.068289.

All of these results hold everything else constant.

**c)**

reg course_eval intro onecredit female minority nnenglish, robust

predict y_resid, resid

reg beauty intro onecredit female minority nnenglish, robust

Name: Julian Bertalli
SID: 3033589368
Section ID: 103
jjbertalli@berkeley.edu

Name: Ben Chu
SID: 3032040857
Section ID: 110
ben-chu@berkeley.edu

predict x_resid, resid

reg y_resid x_resid, robust

```
. reg y_resid x_resid, robust

Linear regression                               Number of obs     =        463
                                                F(1, 461)         =      27.82
                                                Prob > F          =     0.0000
                                                R-squared         =     0.0599
                                                Root MSE          =     .51071

------------------------------------------------------------------------------
             |               Robust
     y_resid |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     x_resid |     .16561   .0313969     5.27   0.000     .1039112    .2273087
       _cons |  -2.82e-09   .0237348    -0.00   1.000    -.0466419    .0466419
------------------------------------------------------------------------------
```

According to the *Frisch–Waugh theorem*, the coefficient on beauty for the multiple regression model in b) can be estimated by the following steps:

1) Regress course_eval on the rest of the regressor variables (call for robust) then get the residuals

2) Regress beauty on the rest of the covariates (call for robust) then get the residuals

3) Run a bivariate regression on the residuals from both step 1) and step 2)

This final regression will yield residuals without the influence of any other outside covariates. This will also isolate the marginal effect; the influence that beauty has on course_eval.

Name: Julian Bertalli
SID: 3033589368
Section ID: 103
jjbertalli@berkeley.edu

Name: Ben Chu
SID: 3032040857
Section ID: 110
ben-chu@berkeley.edu

After completing this process, the new linear regression yields a coefficient estimate of .16561, which is the same estimate coefficient for beauty as obtained in part b). The three step process yields the same estimated coefficient for beauty as that obtained in part b).
.16561 = .16561
Verified

**d)**

| | |
|---|---|
| beauty = 0 | Average beauty |
| intro = 0 | Upper-division |
| onecredit = 0 | Three-credit course |
| female = 0 | Male |
| minority =1 | Black |
| nnenglish =0 | Native English Speaker |

_cons = 4.068289
minority = -.1666154

Beauty has a mean of zero, so if Professor Smith is "average," beauty, he must also have the value of zero for this variable. Furthermore, in the context of Professor Smith's demographics, all of the dummy variables are set to zero aside from the constant and minority (black).

This means that the predicted score for Professor Smith is 4.068289 - .1666154 = 3.9016736

**Problem 3. Education and distance to college**

**a)**

I expect a negative sign for the relationship between the number of completed years of education for young adults and the distance from each student's high school to the nearest 4-year college. It

Name: Julian Bertalli
SID: 3033589368
Section ID: 103
jjbertalli@berkeley.edu

Name: Ben Chu
SID: 3032040857
Section ID: 110
ben-chu@berkeley.edu

makes intuitive sense that being closer to a college will encourage students to attend, which will increase their years of education. Closer college proximity means that the overall cost of education will be lower, which is important since the cost of education is a huge factor when considering if continuing education is worth it, or even feasible. This means that students who live closer to a 4-year college should on average complete more years of education (higher educational attainment). This can be summarized by a negative relationship because on average, smaller distance = higher education. Conversely, higher distance = lower education. The mechanism that I will use to show a negative relationship is regression analysis with Stata.

**b)**

regress yrsed dist, robust

```
. regress yrsed dist, robust

Linear regression                               Number of obs     =      3,796
                                                F(1, 3794)        =      29.83
                                                Prob > F          =     0.0000
                                                R-squared         =     0.0074
                                                Root MSE          =     1.8074
```

| yrsed | Coef. | Robust Std. Err. | t | P>\|t\| | [95% Conf. | Interval] |
|---|---|---|---|---|---|---|
| dist | -.0733727 | .0134334 | -5.46 | 0.000 | -.0997101 | -.0470353 |
| _cons | 13.95586 | .0378112 | 369.09 | 0.000 | 13.88172 | 14.02999 |

Estimated slope: -.0733727

The slope is statistically significant at the 1% significance level since the p-value is so low (0.000). This also means that the results are significant at the 5% and 10% significance levels. Low p-value = high significance.

Name: Julian Bertalli
SID: 3033589368
Section ID: 103
jjbertalli@berkeley.edu

Name: Ben Chu
SID: 3032040857
Section ID: 110
ben-chu@berkeley.edu

Based on the small $R^2$ of 0.0074, the distance from each student's high school to the nearest 4-year college explains a very small fraction of the variance in a student's overall educational attainment measured in years across individuals.

**c)**

reg yrsed dist bytest female black hispanic incomehi ownhome dadcoll momcoll cue80 stwmfg80, robust

```
. reg yrsed dist bytest female black hispanic incomehi ownhome dadcoll momcoll cue80 stwmfg80, robust

Linear regression                               Number of obs     =      3,796
                                                F(11, 3784)       =     183.54
                                                Prob > F          =     0.0000
                                                R-squared         =     0.2829
                                                Root MSE          =     1.5383
```

| yrsed | Coef. | Robust Std. Err. | t | P>\|t\| | [95% Conf. | Interval] |
|---|---|---|---|---|---|---|
| dist | -.0308039 | .0116178 | -2.65 | 0.008 | -.0535816 | -.0080262 |
| bytest | .0924474 | .0030009 | 30.81 | 0.000 | .0865638 | .0983309 |
| female | .1433777 | .0502841 | 2.85 | 0.004 | .0447912 | .2419642 |
| black | .3538083 | .0674994 | 5.24 | 0.000 | .2214695 | .4861471 |
| hispanic | .4023514 | .0737302 | 5.46 | 0.000 | .2577966 | .5469063 |
| incomehi | .3665952 | .0622404 | 5.89 | 0.000 | .2445672 | .4886233 |
| ownhome | .1456416 | .0648174 | 2.25 | 0.025 | .0185612 | .2727221 |
| dadcoll | .5699153 | .0762509 | 7.47 | 0.000 | .4204185 | .7194121 |
| momcoll | .3791836 | .0835917 | 4.54 | 0.000 | .2152945 | .5430728 |
| cue80 | .024418 | .0092692 | 2.63 | 0.008 | .0062449 | .0425911 |
| stwmfg80 | -.0502044 | .0195902 | -2.56 | 0.010 | -.0886128 | -.011796 |
| _cons | 8.861373 | .2410771 | 36.76 | 0.000 | 8.38872 | 9.334027 |

Estimated effect of dist on yrsed: -.0308039

The estimated effect of dist on yrsed is smaller when these characteristics are controlled for. The estimated effect is now -.0308039 (smaller absolute), while before, the effect was -.0733727 (larger absolute). Smaller effect (absolute). This is substantially different from the regression results in part b).

Name: Julian Bertalli
SID: 3033589368
Section ID: 103
jjbertalli@berkeley.edu

Name: Ben Chu
SID: 3032040857
Section ID: 110
ben-chu@berkeley.edu

With a p-value of .008, the value is still statistically significant at the 1% significance level (as well as the 5% and 10% levels).

The new $R^2$ of .2829 is much larger than the $R^2$ of 0.0074 from part b) and the coefficient has dropped significantly in comparison to part b). Without the controls for the characteristics listed, the distribution from part b) was likely being impacted by the impact of several other variables, which means that the results from the regression in part b) did seem to suffer from important omitted variable bias.

**d)**

1 = Father is a College Graduate
0 = Father is not a College Graduate

The positive coefficient dadcoll measures the impact that a student's father graduating college has on the college attainment (in years) of a student, with a binary distribution of 1 = father graduated college or 0 = father did not graduate from college. This measurement is holding all other variables constant to derive the effect of a father's college attendance, while excluding the effects of other variables. According to this sample regression, on average, students with a father who did graduate college (dadcoll = 1) complete .5699153 more years of college than a student with a father who did not graduate college (dadcoll = 0), hence, the positive coefficient value. A student with a father who graduated from college may have higher educational attainment in years than a student with a father who did not graduate from college.

**e)**

cue80 = County Unemployment rate in 1980
stwmfg80 = State Hourly Wage in Manufacturing in 1980

It is important that the terms cue80 and stwmfg80 appear in the regression since they capture the opportunity cost of a person going (or not goin) to college within the context of this problem.

Name: Julian Bertalli
SID: 3033589368
Section ID: 103
jjbertalli@berkeley.edu

Name: Ben Chu
SID: 3032040857
Section ID: 110
ben-chu@berkeley.edu

This means that in this county/state in 1980, the opportunity cost for attending college is the cost of giving up working during the time that will be taken up by attending college, so the wage and unemployment rate plays a big role in the decision to attend college (due to high unemployment or low wage) or go into the workforce and forego college (due to low unemployment and high wage). As the hourly wage for manufacturing jobs in 1980 (stwmfg80) increases, the wages that a person would be giving up in order to attend college increases, which means that the years of education will likely decrease on average since people would rather work and collect a higher wage by working instead of giving up the wage to attend college. As the county's unemployment rate (cue80) rises, fewer people will be employed, meaning that the supply of jobs decreases, meaning that jobs will become more scarce, and therefore, tougher to get. This will make the opportunity cost of attending college lower since the opportunity to even get a job is much tougher, which means that on average, the educational attainment in years should increase. Additionally, being unable to find a job will force some people to have the extra time to go back to school in some capacity.

**f)**

Black male
20 miles from nearest 4-year college
bytest score = 58
1980 family income = $26,000
Owned a home
Mother attended college
Father did not attend college
Unemployment rate = 7.5%
Hourly wage = $9.75
Bob's years of completed schooling using regression in c).

-.0308039(2) + .0924474(58) + .1433777(0) + .3538083(1) + .4023514(0) + .3665952(1) + .1456416(1) + .3791836(1) + .5699153(0) + .024418(7.5) - .0502044(9.75) + 8.861373 = 15.1005852

Name: Julian Bertalli
SID: 3033589368
Section ID: 103
jjbertalli@berkeley.edu

Name: Ben Chu
SID: 3032040857
Section ID: 110
ben-chu@berkeley.edu

(yrsed dist coefficient)(20 miles = 2) + (bytest coefficient)(bytest score = 58) + (if female)(0) + (if black)(1) + (if hispanic)(0) + (if income >$25K)(1) + (if ownhome)(1) + (if momcoll)(1) + (if dadcoll)(0) + (unemployment rate)(rate %) + (hourly wage coefficient)(wage) + constant = prediction of years of schooling

According to the formula, Bob's predicted years of completed schooling is 15.1005852 years.

**Problem 4. The sheepskin effect**

**a)**

I believe that Jaeger and Page estimate their model using only data on white men in this table because they are trying to isolate the "sheepskin effect"/"diploma effect" from any instances of race or gender from confounding the results. By using only one gender and one race, the results are less likely to be attributed to anything besides the "sheepskin effect"/"diploma effect." Using multiple races or genders (or both) would make it difficult to distinguish the biasedness of results within the sample since the returns for education may vary to different degrees based on education for other races or education for other genders.

**b)**

Some of the coefficients included are not reported. The footnote says: "Model also includes Potential Experience and Potential Experience Squared as covariates. Columns (3) and (4) also include dummy variables for zero through eight completed years of education." 12 is the omitted category and the coefficient for 9 is -0.227 (0.049).

Potential Experience = number of years since leaving school
Potential Experience = age - 6 (first year of school) - education (years)
The omitted category, 12, is the referecnce or "baseline" group

Name: Julian Bertalli
SID: 3033589368
Section ID: 103
jjbertalli@berkeley.edu

Name: Ben Chu
SID: 3032040857
Section ID: 110
ben-chu@berkeley.edu

Interpretation of category 9: On average, a person who has completed 9 years of education will earn -0.227 less in log hourly wages compared to a person with the reference group 12 years of education. On average, a person who has completed 9 years of education will earn roughly 22.7 % less wage per hour  (given by the negative sign) compared to a person who has completed 12 years of education. This assumes that both people have the same Potential Experience. It is worth noting that these calculations hold other factors, including Potential Experience, constant.

**c)**

I believe that the effect of the 14th year of education is larger than that of the 15th because the 14th year may represent the attainment of an associate's degree or graduation from a community college (2 year degree). This is assuming that the first 12 years of education represent the completion of high school and years 12 and 14 represent the first two years of college completion. The completion of a 14th year of education shows that the value of a 2-year associate's degree for earnings-attainment in comparison to just a high school diploma. Furthermore, a graduate of a community college (associate's degree) has more value on the job market than a 4-year college dropout during the 15th year of education within the sample. Essentially, a 2-year graduate may be better off than a 4-year dropout with three years of college education, despite the fact that they have a higher amount of education. This may be because failure to complete college turn off employers. Completing only a two-year degree may not be as prestigious or rigorous as a 4-year degree; however, it shows that the candidate has the ability to complete a program. Employers may be biased towards 4-year dropouts in comparison to 2-year graduates.

**d)**

The average difference for the wage a senior could get now vs. the wage they would get after graduation:

**15 years education:** .052 for 15 years of education (currently senior) + .083 for diploma effect (some college) = .052 + .083 = .135 more on average log hourly wages compared to a person with 12 years of education and the same level of experience (high school graducate)

Name: Julian Bertalli
SID: 3033589368
Section ID: 103
jjbertalli@berkeley.edu

Name: Ben Chu
SID: 3032040857
Section ID: 110
ben-chu@berkeley.edu

**16 years education:** .178 for 16 years of education (following graduation) + .245 for earning a Bachelor's degree = .178 + .245 = .423 more on average log hourly wages compared to a person with 12 years of education and the same level of experience (high school graduate)

Average difference in wage: .423 - .135 = .288(100%) = 28.8%
28.8% hourly wage difference

**e)**

Based on column 4, I would rather choose to pursue a professional degree. The marginal effect of getting a professional degree over a bachelor's degree is .286 based on average log earnings per hour. Conversely, the marginal effect of a doctoral degree is only .067. Since .067<.286, a professional degree is the better choice compared to a PhD.

**f)**

Run hypothesis test for the presence of diploma effect:

Null Hypothesis:

$H_0$: The "diploma effects" in column (4) = 0
$\beta_{\text{HighSchool}} = 0$
$\beta_{\text{SomeCollege}} = 0$
$\beta_{\text{OccupationalAssociate's}} = 0$
$\beta_{\text{AcademicAssociate's}} = 0$
$\beta_{\text{Bachelor's}} = 0$
$\beta_{\text{Master's}} = 0$
$\beta_{\text{Professional}} = 0$
$\beta_{\text{Doctoral}} = 0$

Alternative Hypothesis:

Name: Julian Bertalli
SID: 3033589368
Section ID: 103
jjbertalli@berkeley.edu

Name: Ben Chu
SID: 3032040857
Section ID: 110
ben-chu@berkeley.edu

$H_1$: At least one of the "diploma effects" in column (4) does not = 0

28 explanatory variables
- 8 "diploma effect" dummy variables
- 9 zero to eight years of education dummy variables
- 9 nine to eleven and thirteen to eighteen+ years of education dummy variables (exclude 12)
- 2 potential experience and potential experience squared

Assume model includes intercept, which is why -1 is in the formula

Homoskedasticity-only F-statistic:

$$\text{F-statistic} = \frac{(R^2 \text{ unrestricated} - R^2 \text{ restricated}) / q}{(1 - R^2 \text{ unrestricated})/(n - k \text{ unrestricted} - 1)}$$

$$= \frac{(.154 - .147) / 8}{(1 - .154)/(8957 - 28 - 1)} = (.000875/.000094758) = 9.23404884 = 9.234$$

F-statistic = 9.234

5% significance level of $F_{8,\infty}$ critical value = 1.93

Based on the critical value of 1.93, we can reject the null hypothesis: reject that the "diploma effects" in column (4) = 0. This means that there is some evidence in favor of the "diploma effect."

Sample size = 8,957