

HW2 [OLS]: Due Saturday after break week

Question 1

```
In [7]: result = sm.OLS(data_w_dummies['lnepratio'], sm.add_constant(data_w_dummies[['lnooperatingmargin', 'Agriculture Forestry And Fishi
...: 'Services', 'Transportation Communications Electric Gas And Sanitary Service']]), missing='drop').fit()
...: result.summary()
Out[7]:
<class 'statsmodels.iolib.summary.Summary'>
"""
                        OLS Regression Results
=====
Dep. Variable:          lnepratio      R-squared:                0.085
Model:                  OLS          Adj. R-squared:            0.084
Method:                 Least Squares   F-statistic:             1188.
Date:                   Sat, 10 Oct 2020  Prob (F-statistic):       0.00
Time:                   17:09:31       Log-Likelihood:        -1.5595e+05
No. Observations:      115861         AIC:                   3.119e+05
Df Residuals:          115851         BIC:                   3.120e+05
Df Model:               9
Covariance Type:       nonrobust
=====
                        coef      std err      t      P>|t|      [0.025      0.975]
-----
const                    -3.6332      0.009    -419.187    0.000    -3.650    -3.616
lnooperatingmargin        0.2969      0.003     96.466    0.000     0.291     0.303
Agriculture Forestry And Fishing  0.6587      0.058    11.417    0.000     0.546     0.772
Construction              0.3404      0.025    13.352    0.000     0.290     0.390
Finance Insurance And Real Estate -0.1784      0.008   -21.688    0.000    -0.195    -0.162
Manufacturing             -0.0182      0.008    -2.360    0.018    -0.033    -0.003
Mining                    -0.0450      0.018    -2.489    0.013    -0.080    -0.010
Retail Trade              0.1256      0.014     9.187    0.000     0.099     0.152
Services                  -0.2630      0.010   -26.005    0.000    -0.283    -0.243
Transportation Communications Electric Gas And Sanitary Service -0.1285      0.012   -10.640    0.000    -0.152    -0.105
=====
Omnibus:                18757.262   Durbin-Watson:           1.031
Prob(Omnibus):           0.000   Jarque-Bera (JB):       159948.347
Skew:                    0.533   Prob(JB):                0.00
Kurtosis:                8.657   Cond. No.                51.5
=====

Warnings:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
"""
```

Regression output above is taken from Class 4 slides.

Recapping, this is generated in Python, by regressing $\ln(\text{EPS}/\text{price})$ against $\ln(\text{Operating Margin})$ and industry dummy variables. Dataset is panel data on US publicly listed firms. Timeframe for the dataset is approximately 10 years, such that each firm has approximately 10 years of data (unless it is listed or delisted part way through the panel). Each industry dummy variable is 1 if that observation involves a firm belonging to that industry, and 0 otherwise. One of the industry dummies (involving the Wholesale Trade industry) is excluded as the control group.

In the regression output, estimated coefficient on “const” denotes the y-intercept of the regression, estimated coefficient on “lnooperatingmargin” is the slope coefficient on $\ln(\text{Operating Margin})$, and the estimated coefficients on each of the other variables with industry names are estimated coefficients on the industry dummies.

Answer the following questions. Correct answers to all questions are real numbers which you can deduce from the regression output.

- If $\ln(\text{Operating Margin})$ is 0 [hypothetically speaking, this is close to impossible in practice], what is predicted $\ln(\text{EPS}/\text{price})$ for a firm in the wholesale trade industry?
- If $\ln(\text{Operating Margin})$ is 0, what is predicted $\ln(\text{EPS}/\text{price})$ for a firm in the services industry?
- What is the effective “y-intercept” for a firm in the Services industry? [hint: if you got the correct answer for part b, then part c should be extremely simple ...]
- What is the effective “y-intercept” for a firm in the Mining industry?

Question 2

```

=====
                        OLS Regression Results
=====
Dep. Variable:          lnopratio    R-squared:                0.053
Model:                  OLS          Adj. R-squared:            0.052
Method:                 Least Squares   F-statistic:              22.55
Date:                   Fri, 16 Oct 2020   Prob (F-statistic):       2.45e-61
Time:                   20:01:30         Log-Likelihood:           -41725.
No. Observations:      31998          AIC:                     8.348e+04
DF Residuals:          31982          BIC:                     8.361e+04
DF Model:              15
Covariance Type:       cluster

=====
                                coef    std err          z      P>|z|      [0.025    0.975]
-----
const                        -3.6108      0.048     -75.842    0.000     -3.704     -3.518
lnoperatingmargin_Construction  0.2473      0.109      2.266    0.023      0.033      0.461
Construction                 0.7646      0.378      2.021    0.043      0.023      1.506
lnoperatingmargin_Finance Insurance And Real Estate -0.1840      0.194     -0.947    0.343     -0.565      0.197
Finance Insurance And Real Estate -0.8341      0.417     -2.002    0.045     -1.651     -0.018
lnoperatingmargin_Manufacturing -0.8364      0.066     -12.555    0.000     -0.965     -0.707
Manufacturing                 0.2274      0.106      2.151    0.031     -0.035      0.488
lnoperatingmargin_Mining       0.3526      0.084      4.213    0.000      0.189      0.517
Mining                        0.4767      0.143      3.334    0.001      0.196      0.757
lnoperatingmargin_Retail Trade -0.0027      0.082     -1.132    0.258     -0.163      0.158
Retail Trade                  -0.2019      0.218     -0.926    0.354     -0.629      0.225
lnoperatingmargin_Services     0.1619      0.060      2.720    0.007      0.045      0.279
Services                      -0.0084      0.132     -0.063    0.950     -0.268      0.251
lnoperatingmargin_Transportation Communications Electric Gas And Sanitary Service 0.0424      0.075      0.567    0.571     -0.104      0.189
Transportation Communications Electric Gas And Sanitary Service -0.0560      0.163     -0.343    0.732     -0.376      0.264
lnoperatingmargin              0.2399      0.021     11.437    0.000      0.199      0.281

=====
Omnibus:                 8577.755   Durbin-Watson:           1.016
Prob(Omnibus):           0.000   Jarque-Bera (JB):       89878.807
Skew:                    0.988   Prob(JB):               0.00
Kurtosis:                10.969   Cond. No.               144.

=====
Warnings:
[1] Standard Errors are robust to cluster correlation (cluster)
=====

```

This OLS regression output is taken from the Class 4 slides. Dataset structure is the same as for Question 1. However, for this regression, all RHS variables from Q1 are included (with the same meaning), and also, there are additional variables on the RHS, which are interaction variables. Each interaction variable is the numerical value of Ln(Operating Margin) multiplied by an industry dummy variable. This is indicated by the name of the interaction variable. For instance, the interaction variable *lnoperatingmargin_manufacturing* is Ln(Operating Margin)*Manufacturing dummy variable.

Answer the following questions. Similar to question 1, correct answers to all questions are real numbers which you can deduce from the regression output.

- For a 1 unit increase in *lnoperatingmargin*, by how much does predicted Ln(E/P ratio) change by for a firm in the omitted control group (Wholesale Trade industry)
- For a 1 unit increase in *lnoperatingmargin*, by how much does predicted Ln(E/P ratio) change by for a firm in the Mining industry?
- What is the effective slope coefficient for a firm in the mining industry? [hint: once again, if you got part b correct, part c is extremely simple]
- Assume that *lnoperatingmargin* = 1. What is predicted Ln(E/P ratio) for a firm in the Retail Trade industry? [hint: for this part, recall that in this regression, all RHS variables from Q1 are included, with the same meaning as in Q1, in addition to new variables]