Assignment C-1

Estimate the determinants of the change in total employment in US counties over 2010-2015. Your sample consists of [one hundred counties](http://milesfinney.net/491/man_assign/data_I.txt).

1. First, run four regressions: the log/log and linear version of each model below.

*A.* $\hat{y}= b\_{0}+b\_{1}x\_{1}+b\_{2}x\_{2}+b\_{3}x\_{3}+b\_{4}x\_{4}$

*B.* $\hat{y}= b\_{0}+b\_{1}x\_{1}+b\_{2}x\_{2}+b\_{3}x\_{3}+b\_{4}x\_{4}+b\_{5}x\_{5}+b\_{6}x\_{6}+b\_{7}x\_{7}$

**Dependent variables:**

The dependent variable is county growth in total employment from 2010 to 2015. For the log/log model the dependent variable is the natural log of$ \frac{2015 total employment}{2010 total employment}$. In the linear specification, the dependent variable is $\frac{2015 total emp.-2010 total emp.}{2010 total emp. }$. Use the [County Business Patterns](https://www.census.gov/programs-surveys/cbp/data/tables.All.html) to construct the dependent variable.

**Independent variables:**

**x1** - percent of the state’s population over 50 years old in 2010

**x2 = 1** if the EPA designated the county as non-attainment for ground level ozone (eight hour standard) in 2010

 **= 0** otherwise

**x3** - proportion of the county’s residents over 25 who have attained a high school degree or less in 2010

**x4 -** total county population in 2010 (input in 1000’s)

**x5, x6,x7** represent dummy variables indicating the [Census Region](http://www2.census.gov/geo/pdfs/maps-data/maps/reference/us_regdiv.pdf) the county is in

The above variables can be constructed from data found in the [US Census](http://factfinder2.census.gov/faces/nav/jsf/pages/searchresults.xhtml?refresh=t), [EPA](http://www.epa.gov/airquality/greenbook/) and the [Bureau of Labor Statistics](http://www.bls.gov/home.htm).

2. Turn in your SAS regression output as well as the constructed variables used to run the regressions. Also turn in your SAS programs.

3. Write a short paper interpreting the results, addressing the following points.

A. Justify the models. What effect do you expect each of the independent variables to have on the dependent variable? (Your expectation may differ from the estimated relationship).

B. Discuss your results. Do your results correspond with the expectations discussed in part A? Which results are statistically significant? Are your findings sensitive to model specification (log/log or linear)?

C. Which result do you find most surprising? Explain.

D. Does the introduction of the regional dummy variables change the results for the remaining independent variables? Discuss any substantial changes. Interpret the regional dummies.