<div align="center">

Econometrics
Exercise: Panel Data

</div>

NOTE: You will need the package 'plm' for this exercise. So please **before class** open up RStudio, go to the top panel and select Tools, then select Install Packages, then type in plm and hit Install.

Load the data set 'guns_data.csv'.

Save all your code for your regressions in a 'script' file.

1. First lets looked at some 'pooled' regressions.

a. Estimate $log(vio) = \beta_0 + \beta_1 shall + u$

b. Estimate $log(vio) = \beta_0 + \beta_1 shall + \beta_2 incarc\_rate + \beta_3 density + \beta_4 avginc + \beta_5 pop + \beta_6 pb1064 + \beta_6 pw1064 + \beta_7 pm1029 + u$

Interpret the coefficients on 'shall' in the two regressions (careful!! That is a dummy in a log-linear set-up). What do they mean in 'real' terms.

Note that the above pooled regressions completely ignore the fact that we have data on the same states many times. So lets take this into account using a panel data 'random effects' model.

The easiest way to do this is to use a special package designed for panel data. Now hopefully you already downloaded the package. Now to use it just type in:

library(plm) and hit enter

To use this is almost exactly as basic regression with the 'lm' command. The difference is you need to tell it the 'unit' and the 'time' variables and you need to tell it what type of model.

Ex:
mypanel.r $<-$ plm(y~x1+x2,data = mydata, index = c("stateid", "year"), model = "random")

Because 'stateid' is the variable with the id's for the units (States) and 'years' is the variable that has the time index. And "random" tell it to do a random effect model.

2. Now interpret the coefficient on 'shall'. What do they mean in 'real' terms.

So the 'random effect' model recognizes that I have many observations for each unit. But it treats the unit-specific intercept as independent of the Xs - this does not allow us to control for OVB like the 'fixed effect' model we went over in class does.

Lets assume the individual intercepts (the unobserved characteristics) are correlated with our xs and run a FE model:
The only thing you change is: model = "within"

3. Now is the coefficient on 'shall' significant?

Very rarely will the assumptions of the pooled or random effect model hold. There are tests, but we won't get into those. But unless you have very strong reasons to think they do, the default for panel data sets should be fixed effect models.

4. Repeat the above for $ln(rob)$.

Do you think concealed guns limit violent crimes based on this data? Also note the authors of the paper from which this data came do a much more robust investigation of the data than we are doing here.