# ECON 407: Problem Set 3 2/23/2014

This problem set is due by Monday March 10 5pm. Please

1. Turn this in to the Departmental Assistant (Kathy O'Brian) in Morton 102. She will then give you an answer key which will be helpful for studying for the mid-term on the next day. If you turn it in earlier (e.g. this week), you can get the problem set key earlier as well.
2. Turn in a hard copy printout of your responses to these questions in a word (or some other word processing) document. Please incorporate Stata commands, results, and narrative in your written responses to this Problem Set. Combine these into a continuous narrative so I don't have to jump around looking at code and reading results.

## 1 Panel Data Analysis [60 points]

This section of the problem set focuses on the data used in Card and Kreuger's article, "Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania". The American Economic Review 84.4 (Sept. 1994) 772-793. The article can be found here and the codebook for the dataset is here. As described in the codebook, we'll use two versions of the dataset that contains the same information, but is shaped differently. These can be accessed using `webuse Card_Kreuger_wide` and `webuse Card_Kreuger_long`. Remember, to run `webuse set "http://rlhick.people.wm.edu/econ407/data"` before trying to webuse the data.

1. Carefully describe the minimum wage policy being considered in this paper and how the policy is captured in the data.
2. Look carefully at the structure of both datasets and estimate a pooled OLS regression model using the version of the dataset you feel is most appropriate. Describe the specification you choose and take care to specify a model that will enable the quantification for how minimum wage policies change employment. Justify your specification and be clear about your prior hypothesis on the sign of the coefficients. Report on the employment elasticity with respect to the change in the minimum wage policy.
3. Next, estimate a random effects model. Report on the elasticities from this model. Test the random effects model versus the pooled OLS model. State clearly the NULL hypothesis and make a recommendation on which model is appropriate.
4. Estimate a fixed effects model using the 3 approaches we discussed in class: (1) the dummy variable estimator, (2) the de-meaning estimator (this is the default approach using the xtreg command with the fe option), and (3) the first-differences estimator for the same model specification. Compare the coefficients and standard errors from these three approaches to the fixed effects model. Report on the employment elasticity with respect to the change in the minimum wage policy. Note: depending on the model you estimate, you will need to use (or at least find it easier to use) either the "long" or the "wide" version of the data.
5. Given the various models you estimated, what is your overall conclusion about the effect of minimum wage policies based on this data?

## 2 Econometric Theory [40 points]

6. Recall that in a random effects setting, we assumed that the error term for an individual in the model, $v_{it}$ is comprised of a time variant individual-specific effect ($c_i$) and an i.i.d error term ($\epsilon_{it}$). The error in this problem is $v_{it} = c_i + \epsilon_{it}$ The estimating equation in this case is

$$y_{it} = x_{it}\beta + c_i + \epsilon_{it} \tag{1}$$

Now suppose that you collect data on siblings (denoted by $i$) within families (denoted by $j$) and you do this for a number of time periods (denoted by $t$). Now, you believe that the individual error term is comprised of three items: (1) the time invariant unobserved effects as before ($c_i$), (2) a family unobserved time invariant effect ($f_j$), and (3) the i.i.d error term ($\epsilon_{ijt}$). The error in this model is $v_{ijt} = c_i + f_j + \epsilon_{it}$ Let the estimating equation be

$$y_{ijt} = \mathbf{x}_{ijt}\beta + c_i + f_j + \epsilon_{ijt} \tag{2}$$

   (a) What are the implication for strict exogeneity in this problem? Hint: consider our discussion of this issue in the simple panel case and try to extend it to the case where there is an additional unobserved, family specific effect.

   (b) In the random effects framework, we tackled this problem by "fixing" the error structure in the model by way of the matrices $\Sigma$ (for the individual) and $\Omega$ (across the sample). Try to derive an analogous version of $\Omega$ for this problem. FYI: When you ask stata to cluster your standard errors, this question will help you understand what it is doing.

7. Using the fixed effects method of your choosing, show for the general case why time invariant independent variables can not be included in the model.