**Assignment 2: Due Monday, Apr 24 before Midnight**

**Directions:** Respond appropriately to the following questions. Upload your final assignment onto e-learning by the due date. Insert all tables and images, if any, into your word file (or pdf) so that answers are all in one place. Copy and paste your STATA code at the end of the assignment or submit a separate do file. Part of your grade (7 points) will be based on the code, and the remaining will be based on your ability to follow directions and *fully* explain econometric models (73 points). This is an individual assignment. You must turn in your own word document (or pdf). Late submissions within 24 hours will receive 50% of the original points, late submissions within 48 hours will receive 25%, and so on.

1. (Lecture 6) (14 points) Use the data in APPLE to answer this question.
2. (2 point) Define a binary variable as *ecobuy* = 1 if *ecolbs* > 0 and *ecobuy* = 0 if *ecolbs* = 0. In other words, *ecobuy* indicates whether, at the prices given, a family would buy any ecologically friendly apples. What fraction of families claim they would buy ecolabeled apples?
3. (3 points) Estimate the linear probability model

and report the results in the usual form. Carefully interpret the coefficients on the price variables (*ecoprc* and *regprc*).

1. (4 points) Are the nonprice variables jointly significant in the LPM? (Use the usual *F* statistic, even though it is not valid when there is heteroskedasticity.) Which explanatory variable other than the price variables seems to have the most significant effect on the decision to buy ecolabeled apples? Does this make sense to you?
2. (5 points) In the model from part (ii), replace *faminc* with log(*faminc*). Given the *R*2, which model fits the data better? How many estimated probabilities are negative? How many are bigger than one? Should you be concerned? [Hint: Use command predict y to generate fitted values.]
3. (Lecture 7) (11 points) Use the data in EZANDERS for this exercise. The data are on monthly unemployment claims in Anderson Township in Indiana, from January 1980 through November 1988. In 1984, an enterprise zone (EZ) was located in Anderson (as well as other cities in Indiana). [See Papke (1994) for details.]
4. (6 points) Regress log(*uclms*) on a monthly linear time trend and 11 monthly dummy variables. **Copy and paste the stata code you use to generate the monthly linear time trend *t* here.** [Hint: Use *jan* as the base month for the monthly dummy variables.]

What was the overall trend in unemployment claims over this period? (Interpret the coefficient on the time trend.) Is there evidence of seasonality in unemployment claims?

1. (3 points) Add *ez*, a dummy variable equal to one in the months Anderson had an EZ, to the regression in part (i). Does having the enterprise zone seem to decrease unemployment claims? By how much?
2. (2 points) What assumptions do you need to make to attribute the effect in part (ii) to the creation of an EZ?
3. (Lecture 8) (12 points) Use the data in HSEINV for this exercise.
4. (4 points) Find the first order autocorrelation in log(*invpc*) and log(*price*) respectively. Which of the two series may have a unit root? [Hint: Use the command correlate to find the correlation.]
5. (3 points) Based on your findings in part (i), estimate the equation

and report the results in standard form. Interpret the coefficient and determine whether it is statistically significant.

1. (5 points) Now use as the dependent variable. Re-run the equation and report the results in standard form. How do your results of the coefficient change from part (ii)? Is the time trend still significant? Why or why not?
2. (Lecture 8) (12 points) Recall that in the example of testing Efficient Markets Hypothesis, it may be that the expected value of the return at time *t*, given past returns, is a quadratic function of *returnt-1*.
3. (2 points) To check this possibility, use the data in NYSE to estimate

;

report the results in standard form.

1. (4 points) State and test the null hypothesis that does not depend on *returnt-1*. [Hint: There are two restrictions to test here.] What do you conclude?
2. (4 points) Drop from the model, but add the interaction term . Now test the efficient markets hypothesis. [Hint: stata can create lag (or lead) variables using subscripts conveniently. For example, you can use the command gen return\_2 = return[\_n-2] to create *returnt-2* fast.]
3. (2 points) What do you conclude about predicting weekly stock returns based on past stock returns?
4. (Lecture 9) (11 points) Use the data in KIELMC for this exercise.
5. (3 points) The variable *dist* is the distance from each home to the incinerator site, in feet. Consider the model

If building the incinerator reduces the value of homes closer to the site, what is the sign of δ1? What does it mean if β1 > 0?

1. (3 points) Estimate the model from part (i) and report the results in the usual form. Interpret the coefficient on . What do you conclude?
2. (3 points) Add *age*, *age*2, *rooms*, *baths*, log(*intst*), log(*land*), and log(*area*) to the equation. Now, what do you conclude about the effect of the incinerator on housing values?
3. (2 points) Why is the coefficient on log(*dist*) positive and statistically significant in part (ii) but not in part (iii)? What does this say about the controls used in part (iii)?
4. (Lecture 10) (13 points) Use the data in PHILLIPS for this exercise. As we mentioned in Lecture 7, instead of the static Phillips curve model, we can estimate an expectations-augmented Phillips curve of the form

,

where .

1. (3 points) Estimate this equation by OLS and report the results in the usual form. In estimating this equation by OLS, we assumed that the supply shock, *et*, was uncorrelated with *unemt*. If this is false, what can be said about the OLS estimator of β1?
2. (2 points) Suppose that *et* is unpredictable given all past information: . Explain why this makes *unemt-1* a good IV candidate for *unemt*.
3. (3 points) Does *unemt-1* satisfy the instrument relevance assumption? [Hint: You need to run a regression to answer this question.]
4. (5 points) Estimate the expectations augmented Phillips curve by 2SLS using *unemt-1* as an IV for *unemt*. Report the results in the usual form and compare them with the OLS estimates from (i).