**ECO 5720**

**Problem Set 1**

1. Please provide an example of a causal question that you would like to answer using a simple regression model such as

(i) What do *x* and *y* denote? What unobserved factors are in *u*?

(ii) What does the assumption of zero correlation between *x* and *u* imply in your context? Is this likely to be satisfied?

(iii) Based on your answer to part (ii), is *x* exogenous or endogenous?

(iv) What does the assumption of unbiasedness of the OLS estimator imply? In your context, is the OLS estimator of unbiased?

(v) What does the assumption of homoskedasticity imply in your context? Is this likely to be satisfied?

(vi) In general, if *x* is endogenous, does OLS still minimize the sum of squared residuals?

2. Using data from 1988 for houses sold in Andover, Massachusetts, from Kiel and McClain (1995), the following equation relates housing price (*price*) to the distance from a recently built garbage incinerator (*dist*):

(i) Interpret the coefficient on log(*dist*). Please explain the sign and magnitude in terms of *price* and *dist*.

(ii) What other factors about a house affect its price? Might these be correlated with distance from the incinerator?

3. The data set in CEOSAL2 contains information on chief executive officers for U.S. corporations. The variable *salary* is annual compensation, in thousands of dollars, and *ceoten* is prior number of years as company CEO.

(i) Find the average salary.

(ii) How many CEOs are in their first year as CEO (that is, *ceoten* = 0)?

(iii) Estimate the simple regression model

What is the (approximate) percentage increase in salary given one more year as a CEO?