

**Economics 472**  
**Problem Set #10**

(1) Consider a simple model to estimate the effect of personal computer ownership (denoted “PC”) on college grade point average (GPA):

$$GPA_i = \beta_0 + \beta_1 PC_i + u_i.$$

In this model, PC is a binary variable taking the value 1 if the student owns a computer, and is zero otherwise.

(1a) Why might you be concerned about using OLS as a method to estimate  $\beta_0$  and the “causal” effect of computer use on GPA,  $\beta_1$ ? (*Hint:* Think about how mean-independence may potentially be violated.)

(1b) Suppose that you also had data on the parental income of each student. Someone suggests that you should use the parental income variable as an instrument for computer ownership. What conditions would need to be satisfied in order to obtain consistent parameter estimates when parental education is used as an instrumental variable (IV)? Which of these conditions are likely to be violated or satisfied if this particular instrument is used?

(1c) Now, suppose that incoming students are randomly assigned to residence halls. Some of the residence halls are “older” and do not contain their own computer labs, while others contain large computer labs. Explain how you might use this data to construct an instrumental variables estimator. How important is the *random* assignment to residence halls in this context?

(2) Consider the problem of estimating the “causal” effect of the GPA of one’s roommate (RoomGPA) on one’s own GPA:

$$GPA_i = \beta_0 + \beta_1 RoomGPA_i + v_i.$$

(2a) If this is, in fact, the “production function” for GPA, why does it imply a necessary simultaneity problem?

(2b) What does your answer in (2a) suggest about the performance of the OLS estimator for this model?

(2c) Can you suggest any possible instrumental variables strategies for estimating this model? What type of data would be required to implement your approach?