Solutions: Problem Set #10

(1a) Mean-independence could be violated for a variety of reasons. First, highly-motivated students may be more likely to purchase computers. (They decide to purchase a computer, simply because they think it will give them an "edge," and this motivation also is correlated with GPA). In this case, we have an omitted variables problem leading to a mean-independence violation. Second, students from wealthier families may be more likely to purchase a computer. If wealthier families also tend to invest more in the human capital of their children, and this human capital also helps GPA, then again, we have an omitted variables bias problem and mean-independence violation.

(1b) The second story described in (1) suggests that this may not be a valid instrument. Sure, parental income is probably strongly correlated with PC ownership (instrument relevance), but at the same time, parental income may have a direct effect on GPA independently of its effect on PC ownership. In this case, we would violate instrument exogeneity.

(1c) Construct a dummy variable called z which equals one if the student lives in a hall with a computer lab, and equals zero otherwise. We would then apply the IV estimator:

$$\hat{\beta}_{1,IV} = \frac{\sum_{i=1}^{n}(GPA_i - \overline{GPA})(z_i - \overline{z})}{\sum_{i=1}^{n}(PC_i - \overline{PC})(z_i - \overline{z})}.$$

The instrument is seemingly relevant - students who have access to a computer lab in their residence hall are probably less likely to purchase a PC. At the same time, if there is no lab in the residence hall, students are probably likely to purchase a computer since, if you did not own a PC, you would need to leave the residence hall to do computer-based assignments, check e-mail, etc. (If it is cold, there are non-trivial costs for doing this!)

The random assignment also makes the instrument plausibly exogenous - there is nothing in the model that would generate a correlation between z and GPA. If students get to choose their halls, there is an obvious selection bias problem - it may

be the case that high achieving students, or students of lower incomes will select into the halls with computer labs.

(2)

(2a) If this is the correct model, then it also applies to the GPA of the roommate:

$$RoomGPA_i = \delta_0 + \delta_1 GPA_i + v_i.$$

So, we have a simultaneous equation model, and thus a mean-independence violation.

(2b) From (2a), we violate mean-independence, and thus do not want to use OLS.

(2c) There is no "correct" answer to this problem. What we need is some variable that affects the performance of the roommate without affecting one's own performance (conditioned on the achievement of the roommate). Some possibilities might include if the roommate was or became ill (and hopefully not contagious, as this would clearly not be a valid IV!!), if the roommate experienced a death in the family, etc.