## SIMON FRASER UNIVERSITY

Department of Economics

Econ 835
Prof. Kasa
Quantitative Methods

## FINAL EXAM

Answer the following questions True, False, or Uncertain. Briefly explain your answers. No credit without explanation. (5 points each).

1. The Durbin-Watson statistic is not valid when one of the regressors is a lagged dependent variable.
2. An advantage of Instrumental Variables estimators is that they tend to have lower variances than Ordinary Least Squares estimators.
3. The White test for heteroskedasticity is not valid if the error term is serially correlated.
4. It is better to use Weighted Least Squares than Ordinary Least Squares when you suspect the error term is correlated with one of the regressors.
5. Least Squares is an unbiased estimator even if a model's error term is serially correlated.

The following questions are short answer. Briefly explain your answer. Clarity will be rewarded.
6. (5 points). Suppose you are estimating a model relating the number of campus crimes to student enrollment using a sample of colleges. However, since reporting of crime statistics is voluntary, you don't have data from some of the colleges. Should you be worried about this? If so, which of the basic regression assumptions is violated?
7. (5 points) Suppose you compute both the standard and the heteroskedasticity-robust (White) standard errors and they turn out to be quite different. Which should you use? Why?
8. (10 points). In 1989 there was a major earthquake in San Francisco. As a result, some schools were closed for the following year. The students who attended these schools were assigned to nearby schools. Explain why this event might be useful in studying the effect of class size on school performance. Why might there be problems in just regressing test scores on class size during a "normal" period?
9. (10 points). Consider the regression model: $y_{1}=\beta_{0}+\beta_{1} y_{2}+\beta_{2} x_{1}+u$. Suppose you are concerned that $y_{2}$ is "endogenous" (i.e., correlated with $u$ ). Also suppose that $z$ is a candidate instrument for $y_{2}$. Explain in detail how to do a Hausman Test to check for the endogeneity of $y_{2}$. What is your null and alternative hypothesis?
10. (15 points). Consider a standard wage equation: $\log ($ wage $)=\beta_{0}+\beta_{1} e d u c+\beta_{2}$ exper $+u$.
(a) Explain how and why an excluded variable like 'ability' might bias our estimate of $\beta_{1}$. Are we likely to overestimate or underestimate the rate of return to education?
(b) Briefly discuss the use of a proxy variable to address this problem. Provide an explicit example if you can. Briefly discuss how the use of an instrumental variable could address this problem. What two conditions would a variable have to satisfy to be a valid instrument? Provide an explicit example if you can.
(c) Explain why a proxy for ability is likely to be a poor instrument for education.
11. (15 points). There has been a lot of discussion in the U.S. recently about banning the use of cell phones while driving. One proposal would leave it up to the individual states to decide. Suppose this happens soon, and some states pass laws banning cell phone use and some states do not.
Suppose it's the future, and assume you have survey data from two different years, one before the laws were passed and one after. The survey asks people what state they live in, whether they use a cell phone while driving, and whether they've had any auto accidents during the year.
(a) Explain how you would specify a model to estimate the effectiveness of the law in reducing auto accidents. Be sure to allow for aggregate trends in auto accidents and for potential differences across states in the likelihood of getting in an accident, due perhaps to differences in road congestion or weather conditions.
(b) Now suppose you don't have individual survey data, but instead are given statewide averages. How does this data structure differ from the one in part (a)? What kind of estimation strategy could you now use?
12. (15 points). Suppose you want to estimate the causal effect of skipping classes on final exam scores. A simple regression would be:

$$
\text { score }=\beta_{0}+\beta_{1} \text { skipped }+u
$$

where score is the final exam score and skipped is the total number of lectures missed during the semester. (Suppose either students truthfully report this, or the instructor actually took attendance!)
(a) Is OLS likely to yield reliable estimates of the ceteris paribus effect of skipping classes on exam scores? Why or why not?
(b) Describe how an instrumental variable estimation strategy might be used instead. What conditions would an instrument have to satisfy? Can you think of an example?
(c) Suppose you can't find an instrument. However, you do find data on each person's previous GPA and their SAT scores. Describe how you could use this data to produce more reliable estimates of the effect of skipping lectures.

