Note to students: If you would like to get your exam returned to your mailbox, sign the consent form below. If you would not like your exam returned to your mailbox, leave this form blank.

STUDENT CONSENT FOR RELEASE OF STUDENT INFORMATION

(Buckley Waiver)

I hereby authorize the UCSD Economics Department to return my graded final examination paper by placing it in my mailbox. I understand that the return of my examination paper as described above may result in disclosure of personally identifiable information, that is not public information as defined in UCSD PPM 160-2, and I hereby consent to the disclosure of such information.

Quarter__________Course____________Date___________________
Instructor ____
Print Name____
Signature ____

Econ 220B Final Exam
Winter 2019

DIRECTIONS: No books or notes of any kind are allowed. 250 points are possible on this exam.

1.) (25 points possible) A researcher is studying the Troubled Asset Relief Program of 2008 in which the government purchased certain assets held by banks that the market had come to consider "toxic." A researcher wants to study whether these purchases helped banks to make more new good loans than they otherwise would have. The researcher has data on a cross-section of banks $i = 1, 2, ..., n$ where y_i is the number of new loans extended by bank i in 2009 minus the new loans extended by bank i in 2008, and x_i is the quantity of troubled assets purchased by the government from bank i in the fall of 2008. The researcher estimated the following model by OLS:

$$y_i = \alpha + x_i\beta + u_i \qquad i = 1, ..., n.$$

a.) (10 points) If you could imagine running an ideal controlled experiment to answer the researcher's question, what would it be?

b.) (10 points) Would the OLS estimate of β be biased? Why? Which way would you expect the bias to go?

c.) (5 points) Is there a simple test you could perform on the OLS residuals $\hat{u}_i$ that might provide evidence in support of the answer you gave to part (b)?

2.) (25 points total) Let y_t be a scalar and $\mathbf{x}_t$ and $\mathbf{z}_t$ $(k \times 1)$ vectors where $(y_t, \mathbf{x}_t', \mathbf{z}_t')'$ is strictly stationary and ergodic. Consider the model

$$y_t = \mathbf{z}_t'\boldsymbol{\beta} + u_t.$$

The instrumental variable (IV) estimator of $\boldsymbol{\beta}$ is defined by

$$\hat{\boldsymbol{\beta}}_{IV} = \left(\sum_{t=1}^{T} \mathbf{x}_t\mathbf{z}_t'\right)^{-1} \left(\sum_{t=1}^{T} \mathbf{x}_t y_t\right).$$

a.) (15 points) Calculate the plim of $\hat{\boldsymbol{\beta}}_{IV}$. What additional assumptions did you need beyond those stated above to derive this result?

b.) (10 points) Suppose the researcher is not interested in estimating a causal effect but only wants to use $\mathbf{z}_t$ to predict the value of y_t. Would plim$(\hat{\boldsymbol{\beta}}_{IV})$ be an object the researcher would be interested in? Why?

3.) (40 points total) Consider the regression model:

$$y_t = \mathbf{z}_{t1}'\boldsymbol{\beta}_1 + \mathbf{z}_{t2}'\boldsymbol{\beta}_2 + u_t$$

$$E(\mathbf{z}_{t1}u_t) = \underset{(k_1\times 1)}{\mathbf{0}}$$

$$E(\mathbf{z}_{t2}u_t) = \underset{(k_2\times 1)}{\mathbf{a}}.$$

Note that no element of the vector $\mathbf{a}$ is zero. The researcher has available an $(r \times 1)$ vector of instruments $\mathbf{x}_t$ for which $E(\mathbf{x}_t u_t) = \underset{(r\times 1)}{\mathbf{0}}$.

a.) (5 points total) What does it mean to say that the instruments $\mathbf{x}_t$ are valid?

b.) (5 points total) What does it mean to say that the instruments $\mathbf{x}_t$ are relevant?

c.) (10 points total) What does it mean to say that the instruments $\mathbf{x}_t$ are weak? Describe a simple test for weak instruments based on OLS F-statistics.

d.) (15 points total) Describe a test of the null hypothesis $H_0 : \boldsymbol{\beta}_2 = \mathbf{0}$ that would have good small-sample properties even if instruments are weak. Describe the test in the **simplest** way you can for this particular example rather than as a special case of some more general result.

e.) (5 points total) Under what conditions would the test you proposed in (d) have the exactly correct size in a small sample?

4.) (160 points total) This question asks you to explore a generalized method of moments interpretation of maximum likelihood estimation of the Gaussian regression model. The questions are cumulative, but even if you are not sure of one answer, go ahead and try to answer the next one anyway based on other things you may know about MLE, OLS, or GMM. Even if you're sure you did the previous step correctly, comparing your answer at each step with other things you know is probably a good way to check that you did everything correctly, and certainly is the way to get back on track if you do something wrong.

Gaussian regression model:

$$y_t = \mathbf{x}_t'\boldsymbol{\beta} + u_t$$

$$u_t \sim \text{ i.i.d. } N(0, \sigma^2)$$

u_t is independent of the $(k \times 1)$ vector $\mathbf{x}_s$ for all t and s.

Some facts about the Gaussian regression model that you may find useful:

$$\ell(\boldsymbol{\theta}) = \textstyle\sum_{t=1}^{T} \ell_t(\boldsymbol{\theta})$$

$$\ell_t(\boldsymbol{\theta}) = -(1/2)\log(2\pi) - (1/2)\log(\sigma^2) - \frac{(y_t - \mathbf{x}_t'\boldsymbol{\beta})^2}{2\sigma^2}$$

$$\boldsymbol{\theta} = (\boldsymbol{\beta}', \sigma^2)'$$

$$E(u_t^3) = 0$$

$$E(u_t^2 - \sigma^2)^2 = 2\sigma^4.$$

Generalized method of moments:

$$\exists! \underset{(a\times 1)}{\boldsymbol{\theta}_0} : E(\mathbf{h}(\boldsymbol{\theta}_0, \mathbf{w}_t) = \underset{(r\times 1)}{\mathbf{0}}$$

$$\mathbf{g}(\boldsymbol{\theta}; \mathbf{Y}_T) = T^{-1} \textstyle\sum_{t=1}^{T} \mathbf{h}(\boldsymbol{\theta}, \mathbf{w}_t)$$

$$\hat{\boldsymbol{\theta}}_{GMM} = \arg\min_{\boldsymbol{\theta}} T[\mathbf{g}(\boldsymbol{\theta}; \mathbf{Y}_T)]'\mathbf{S}^{-1}[\mathbf{g}(\boldsymbol{\theta}; \mathbf{Y}_T)]$$

$$\hat{\mathbf{D}}' = \left.\frac{\partial \mathbf{g}(\boldsymbol{\theta}, \mathbf{Y}_T)}{\partial \boldsymbol{\theta}'}\right|_{\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}_{GMM}}$$

$$\mathbf{D}' = \text{plim}(\hat{\mathbf{D}}')$$

$$\hat{\mathbf{S}} = T^{-1} \textstyle\sum_{t=1}^{T} [\mathbf{h}(\hat{\boldsymbol{\theta}}_{GMM}, \mathbf{w}_t)][\mathbf{h}(\hat{\boldsymbol{\theta}}_{GMM}, \mathbf{w}_t)]'$$

$$\mathbf{S} = \text{plim}(\hat{\mathbf{S}})$$

$$\sqrt{T}(\hat{\boldsymbol{\theta}}_{GMM} - \boldsymbol{\theta}_0) \xrightarrow{L} N(\mathbf{0}, \mathbf{V})$$

$$\mathbf{V} = (\mathbf{D}\mathbf{S}^{-1}\mathbf{D}')^{-1}.$$

a.) (20 points) Calculate the score for observation t associated with the Gaussian regression model and prove that it follows a martingale-difference sequence. Note that for full credit you should base your proof solely on the facts stated above about the Gaussian regression model and not attempt to derive it as a special case of some more general result about maximum likelihood estimation.

b.) (20 points) Calculate the maximum-likelihood estimates $\hat{\boldsymbol{\beta}}_{MLE}$ and $\hat{\sigma}^2_{MLE}$.

c.) (20 points) Use either result (a) or the first-order conditions from (b) (whichever is easiest for you) to show that the MLE can be viewed as a special case of GMM. What are the values of a, r, and $\mathbf{h}(\boldsymbol{\theta}, \mathbf{w}_t)$ for this interpretation?

d.) (20 points) Calculate $\hat{\mathbf{D}}'$ and $\mathbf{D}' = \text{plim}(\hat{\mathbf{D}}')$ for your interpretation in (c).

e.) (20 points) Calculate $\hat{\mathbf{S}}$ and $\mathbf{S} = \text{plim}(\hat{\mathbf{S}})$ for your interpretation in (c). What is the relation between $\mathbf{D}'$ and **S?**

f.) (10 points) Use results (d) and (e) to calculate the value of $\mathbf{V}$. Comment on the relation between this expression and the usual OLS standard errors.

g.) (10 points) Suppose we use $\hat{\mathbf{V}} = (\hat{\mathbf{D}}\hat{\mathbf{S}}^{-1}\hat{\mathbf{D}}')^{-1}$ for purposes of calculating standard errors for the OLS estimate of $\boldsymbol{\beta}$. Comment on the relation between this expression and White standard errors.

h.) (20 points) Suppose you want to use the estimate $\hat{\mathbf{V}}$ and the fact that $\sqrt{T}(\hat{\boldsymbol{\theta}}_{GMM} - \boldsymbol{\theta}_0) \xrightarrow{L} N(\mathbf{0}, \mathbf{V})$ to test the null hypothesis that $H_0 : \mathbf{q}(\boldsymbol{\theta}_0) = \underset{(m\times 1)}{\mathbf{0}}$ where $\mathbf{q}(.)$ is a known $(m \times 1)$ function with continuous first derivatives. State the test statistic you would use and its asymptotic distribution. Note that you should be able to answer this question even if you have drawn a blank on all of questions (i) through (g)– the answer should simply be written as a function of $\hat{\boldsymbol{\theta}}$ and $\hat{\mathbf{V}}$. Hint: if you can't answer this, try to answer for the case when $\mathbf{q}(.)$ is a known linear function.

i.) (20 points) Alternatively, suppose you wanted to test the hypothesis H_0 in question (h) using Hansen's J test of the overidentifying restrictions in GMM. State the test statistic you would use and its asymptotic distribution.