Econometrics--Econ 388

Spring 2004, Richard Butler

Final Exam

your name____

Section	Problem	Points Possible
I	1-20	3 points each
II	21	15 points
	22	15 points
	23	5 points
	24	5 points
	25	15 points
	26	5 points
III	27	20 points
IV	28	30 points
	29	30 points

I. Define or explain the following terms:

1. dummy variable-

2. dummy variable trap-

3. conditional distribution of Y given X-

4. logistic regression model-

5. random variable -

6. instrumental variables-

7. structural vs. reduced form equations-

8. asymptotic efficiency-

9. law of large numbers-

10. omitted variable bias-

11. central limit theorem-

12. p-value-

13. maximum likelihood estimation criterion-

14. endogenous vs exogenous variables-

15. Breusch-Pagan test-

16. dynamically complete models-

17. feasible generalized least squares--

18. Hausman test-

19.plim--

20. Durbin-Watson test-

II. Some Concepts

21. A random variable W is known to be distributed either N(1,3)—that is, normal with mean 1 and variance 3—or it is distributed N(2,3). The null hypothesis is that the mean equals 1. The alternative is that the mean equals 2.

a. If we have a sample of size 100, what is the cutoff for a test using the mean with the probability of a type one error of .01?

b. What is the probability of a type II error in this problem?

c. If the alternative was N(3,3) instead of N(2,3), would the probability of a type II error increase or decrease? Why?

22. What do the following Shazam programs do:

a. 2SLS A B C (C D)
2SLS B A D (C D)

b.
```
OLS Y X1 X2 / RESID=E
GENR ESQ = E*E
GENR X1SQ = X1*X1
GENR X1SQ = X2*X2
GENR X1X2 = X1*X2
OLS E2 X1 X2 X1SQ X2SQ X1X2
GEN1 LM = $N*$R2
PRINT LM
```

c. What restrictions are imposed on the (slope) coefficients in this model

$$Y = \alpha_0 + \alpha_1 X_1 + \alpha_2 X_2 + \alpha_3 X_3 + \varepsilon$$

by the following Shazam code:

```
GENR Y_X3 = Y - X3
GENR X1X2 = X1 + X2
OLS Y_X3 X1X2
```

23. Explain why the following is true, false, or sometimes true:
a. "The cross section data we used from Ferret for our papers is not likely to exhibit serial correlation."

24. “Y is orthogonally projected onto the X-space (a space generated by a linear combination of the Xs) by a matrix that is idempotent and symmetric.”

25. “The Butler estimator for the linear model, $Y = X\beta + \mu$, is as follows

Butler estimator $\equiv \breve{\beta} = (X'X)^{-1}X'Y + \frac{3.1417}{n}$ where n=sample size, and X and Y are the matrices of the independent and dependent variables respectively (and the usual assumptions hold). True or False: Butler’s estimator is consistent and unbiased.”

26. Write out the Shazam code to get the predicted values for the linear probability model (OLS with a dummy dependent variable), then appropriately adjust them, form the weights, and do weighted least squares to correct for heteroskedasticity.

III. An Application

27. Suppose that for the standard regression model, $y = X\beta + \mu$, we rescale both the independent variables and the dependent variable by non-zero variables $c_0, c_1, c_2, \ldots, c_k$ by regressing $c_0 y_i$ on a constant $c_1, c_2 x_{2i}, c_3 x_{3i}, \ldots, c_k x_{ki}$ (so there are k regressors, including an intercept which is measured as c_1 instead of 1). In other words, instead of the OLS estimator $\hat{\beta} = (X'X)^{-1}X'Y$, you do OLS on the transformed data $\hat{\beta}^* = (X^{*\prime}X^*)^{-1}X^{*\prime}Y^*$, where the transformed data can be described by:

$$X^* = XC \text{ where } C = \begin{bmatrix} c_1 & 0 & 0 & 0 & 0 \\ 0 & c_2 & 0 & 0 & 0 \\ 0 & 0 & c_3 & 0 & 0 \\ 0 & 0 & 0 & \ldots & 0 \\ 0 & 0 & 0 & 0 & c_k \end{bmatrix} \text{ and } Y^* = C_0 Y \text{ , } C_0 = \begin{bmatrix} c_0 & 0 & 0 & 0 & 0 \\ 0 & c_0 & 0 & 0 & 0 \\ 0 & 0 & c_0 & 0 & 0 \\ 0 & 0 & 0 & \ldots & 0 \\ 0 & 0 & 0 & 0 & c_0 \end{bmatrix} = c_0 I .$$

Prove that the i^{th} beta between the transformed data, and the untransformed data, have the following relationship: $\hat{\beta}_i^* = \frac{c_0}{c_i}\hat{\beta}_i$.

IV. A Few Results

28. Prove that s^2, the estimator of the variance of μ_i (where μ_i is the error term in the classical regression model), is unbiased using matrix algebra.

29. Define the following serial stochastic process and show whether each is covariance stationary and weakly dependent:
a) random walk
b) moving average of order 2
c) first order autoregressive process