Second Exam: Economics 388, Econometrics Spring 2004 in R. Butler's section

YOUR NAME:____

Section I (30 points) Questions 1-10 (3 points each)
Section II (20 points) Questions 11-12 (10 points each)
Section III (50 points) Questions 13-14 (15 points each)
Question 15 (20 points)

Section I. Define or explain the following terms (3 points each)

1. log-likelihood ratio test-

2. beta coefficient-

3. Goldfeld-Quandt test-

4. variance of the prediction error-

5. maximum likelihood estimation--

6. "hetcov" option in Shazam-

7. dummy variable trap --

8. Breusch-Pagan test--

9. asymptotic efficient-

10. "diagnos / het" option in Shazam-

II. Some Fun Stuff: True, False or Uncertain Questions (11-14); you are graded on your explanation and not on whether you guessed T or F correctly.

11. T,F or U: "An estimated age-coefficient value of ".05" in a **binomial logit** (or binary logit, logistic regression, or just logit) indicates that for each additional year of age, the probability of marriage increases by 5 percent."

12. T,F or U: "The central limit theorem says that a sample mean, calculated from a distribution with mean μ and variance σ^2 is normally distributed in every sample."

III. Really Fun Stuff

13. 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 $\hat{\beta}^* = (X^{*}{}'X^*)^{-1}X^{*}{}'Y^*$, where the transformed data can be described by:

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

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

14. Data from GPA2.RAW generated the regression below, where

sat=combined SAT score

hsize=size of the individual's high school graduating class (in hundreds)

hsizesq=square of hsize

female=1 if female, 0 if male

black=1 if black, 0 in non-black

```
R-SQUARE =   0.0832     R-SQUARE ADJUSTED =   0.0823
VARIANCE OF THE ESTIMATE-SIGMA**2 =   17834.
STANDARD ERROR OF THE ESTIMATE-SIGMA =   133.54
SUM OF SQUARED ERRORS-SSE=  0.73689E+08
MEAN OF DEPENDENT VARIABLE =   1030.3
LOG OF THE LIKELIHOOD FUNCTION = -26115.9
```

VARIABLE NAME	ESTIMATED COEFFICIENT	STANDARD ERROR	T-RATIO 4132 DF	P-VALUE	PARTIAL CORR.	STANDARDIZED COEFFICIENT	ELASTICITY AT MEANS
HSIZE	19.115	3.837	4.982	0.000	0.077	0.2381	0.0519
HSIZESQ	-2.1894	0.5278	-4.148	0.000	-0.064	-0.1983	-0.0231
FEMALE	-41.608	4.175	-9.967	0.000	-0.153	-0.1485	-0.0182
BLACK	-139.29	9.097	-15.31	0.000	-0.232	-0.2285	-0.0075
CONSTANT	1027.0	6.290	163.3	0.000	0.930	0.0000	0.9968

a). Should hsizesq be in the regression? In terms of optimal high school size, what does it imply?

b). What do the estimated coefficients on the female and black dummy variables indicate (both in magnitude and statistical importance)?

c). Is this one of those regressions where I should worry about interpreting the constant term? Why or why not?

15. Suppose that for the general linear regression model, $y = X\beta + \mu$, the first four modeling assumptions are satisfied. Prove that the OLS estimator for β, namely $\hat{\beta}$, is a consistent estimator. (You can assume that $\text{plim}\left(\frac{X'X}{n}\right) = \Omega$, a positive definite, symmetric matrix