First Exam: Economics 388, Econometrics Spring 2001 in R. Butler's section
YOUR NAME:_____
Section I (30 points) Questions 1-10 (3 points each)
Section II (40 points) Questions 11-14 (10 points each)
Section III (30 points) Questions

Section I. Define or explain the following terms (3 points each)
1. covariance between two random variables, X and Y-

2. orthogonal projection-

3. cross section data-

4. Gauss-Markov Theorem-

5. null hypothesis vs. altenative hypothesis-

6. symmetric matrix-

7. heteroskedasticity-

8. the *linear* part of linear regression (what is it?)-

9. empirical rule for normal distributions-

10. type I error-

II. Some Concepts

11. The following data gathered from Provo give the joint distribution of f(X,Y) of per capita disposable income X (in thousands of dollars) and savings ratio Y. The cell entries are the proportions for the indicated values of X and Y:

Y \ X=	1.5	3.0	4.5	6.0
.1	.05	.10	.15	.05
0.0	.10	.10	.15	.00
-.1	.15	.05	.05	.05

A. calculate the marginal probability densities f(X) and f(Y)

B. Calculate E(Y) and V(Y) (no credit unless you show the right formulas).

C. Calculate the conditional probability density f(X|Y=0.0) (again, no credit unless you show the right formulas)

D. Are X and Y independent? Why or why not?

12. Prove the following result (assume that linear combinations of normal variables are also normal): "If y is distributed $N(\mu_y,\Sigma_y)$, then $z = Ay$ is distributed $N(\mu_z = A\mu_y;\ \Sigma_z = A\Sigma_yA')$ where A is a matrix of constants."

13. Prove that the least squares estimator for the variance, s^2, is unbiased using matrix algebra.

14. Indicate whether the following statements are True, False or Uncertain (indicate which, you are graded only on your explanation for your answer):

a. . In the sample regression model, $\hat{\beta}$ is a random variable because it is an estimate of the β -vector in the population model, and the β in the population model is random."

b. "In the simple regression model, the variance of the estimated slope coefficient increases as: 1) the sample size gets larger, or 2)the variation in the independent variable, X_i gets larger."

II. 15. An Application--the program and output from Shazam is given below. Use them to answer the questions that follow (on the next page):

```
*accessing wooldridge data on HPRICE1.RAW, p. 106, 160 in Wooldridge text
*  1. price                    house price, in dollars
*  2. assess                   assessed value, in dollars
*  3. bdrms                    number of bedrooms
*  4. lotsize                  size of lot, square feet
*  5. sqrft                    size of house, square feet
*  6. colonial                 =1 if home is colonial style
*  7. lprice                   log(price)
*  8. lassess                  log(assess
*  9. llotsize                 log(lotsize)
* 10. lsqrft                   log(sqrft)

file path i:\my2000docs\byu\econ388\textdata\
sample 1 88
read(hprice1.raw) price assess bdrms lotsize sqrft colonial lprice lassess  llotsize
lsqrft
genr lbdrms=log(bdrms)
ols price bdrms sqrft colonial lotsize
test
      test bdrms=0
      test sqrft=0
end
ols price lbdrms lsqrft colonial llotsize
end
stop

R-SQUARE =   0.6758     R-SQUARE ADJUSTED =   0.6602

                     ANALYSIS OF VARIANCE - FROM MEAN
                       SS          DF             MS                F
 REGRESSION       0.62028E+12      4.       0.15507E+12            43.252
 ERROR            0.29758E+12     83.       0.35853E+10           P-VALUE
 TOTAL            0.91785E+12     87.       0.10550E+11             0.000


 VARIABLE   ESTIMATED  STANDARD   T-RATIO        PARTIAL STANDARDIZED ELASTICITY
   NAME    COEFFICIENT   ERROR      83 DF   P-VALUE CORR. COEFFICIENT  AT MEANS
 BDRMS      11004.      9515.       1.156     0.251 0.126     0.0901     0.1338
 SQRFT      124.24      13.34       9.314     0.000 0.715     0.6981     0.8523
 COLONIAL   13716.     0.1464E+05  0.9370     0.351 0.102     0.0619     0.0324
 LOTSIZE    2.0758     0.6427      3.230     0.002 0.334     0.2056     0.0638
 CONSTANT  -24127.     0.2960E+05 -0.8150     0.417-0.089     0.0000    -0.0822
 |_test
 |_ test bdrms=0
 |_ test sqrft=0
 |_end
 F STATISTIC =   68.812862     WITH    2 AND   83 D.F.  P-VALUE= 0.00000

|_ols price lbdrms lsqrft colonial llotsize


  R-SQUARE =   0.6720     R-SQUARE ADJUSTED =   0.6562
```

```
                    ANALYSIS OF VARIANCE - FROM MEAN
                      SS          DF            MS                F
REGRESSION       0.61682E+12      4.       0.15421E+12            42.517
ERROR            0.30103E+12     83.       0.36269E+10          P-VALUE
TOTAL            0.91785E+12     87.       0.10550E+11            0.000


VARIABLE   ESTIMATED  STANDARD   T-RATIO        PARTIAL STANDARDIZED ELASTICITY
  NAME    COEFFICIENT   ERROR      83 DF   P-VALUE CORR. COEFFICIENT  AT MEANS
LBDRMS      53972.     0.3508E+05   1.539     0.128 0.167     0.1195     0.2291
LSQRFT    0.23181E+06 0.3044E+05   7.615     0.000 0.641     0.5838     5.9800
COLONIAL    5095.8     0.1483E+05  0.3436     0.732 0.038     0.0230     0.0120
LLOTSIZE    62275.     0.1249E+05   4.984     0.000 0.480     0.3299     1.8892
CONSTANT -0.20872E+07 0.2082E+06  -10.03     0.000-0.740     0.0000    -7.1103
|_end
|_stop
```

A. Which model seems to fit better, the one with log-transformations of the independent variables or the one without the transformation? Why?

B. What is being tested with the TEST statement following the first OLS command? What is indicated in the test statistics?

C. Is it important to have a Colonial house? (be as specific as you can in your answer)

D. How much would the value of a house increase, typically, if the homeowners added one more bedroom with 150 square feet (the bedroom was 10' by 15')?