

# Homework

Econometrics

July 24, 1999

## 1 General Instructions

The purpose of these exercises is to help you become more comfortable with the tools of econometrics. To accomplish this, you are to write IML programs to obtain the results for the problems found below. Doing these exercises will contribute to your matrix algebra, computing, and overall econometric skills. The data for these exercises can be found in *Econometric Analysis* (Greene 1997) or at the web site:

<http://www.stern.nyu.edu/~wgreene/datasets.htm>

Most of these exercises can be found in Greene's book which also includes the numerical results for some of the problems found below. Feel free to use the book to provide check figures for key results. An additional source of check figures is the use of the canned SAS procedures AUTOREG and MODEL.

You may turn in your answers in the form of annotated notes made on your SAS output. Otherwise, make sure your results and answers are easy to identify and supported by relevant IML programs and SAS output. In order to obtain credit for the exercises you must turn in both the IML programs and output. Each of the hypothesis tests need to be carried out fully, meaning that you should state the null and alternative hypotheses, the distribution of your test statistic under the null hypothesis, your decision criterion, and your decision. Unless otherwise stated, conduct all hypothesis tests at the 5% level.

The homework is to be air mailed to me no later than August 31. Feel free to ask questions or seek help from me if needed. Good luck.

## Problem 1

Data on gasoline consumption in the United States from years 1960-1982 appear in exercise 6.22 (Greene 1997). The variables are:

- $G$  = total gasoline consumption
- $P_g$  = the price of gasoline
- $Y$  = per capita disposable income
- $P_{nc}$  = price index of new cars
- $P_{uc}$  = price index of use cars
- $P_{pt}$  = price index of public transportation
- $P_d$  = price index of durable goods
- $P_n$  = aggregate price index of nondurable goods
- $P_s$  = aggregate price index of consumer services

$$\ln G = \beta_1 + \beta_{P_g} \ln P_g + \beta_Y \ln Y + \beta_{nc} \ln P_{nc} + \beta_{uc} \ln P_{uc} + \beta_{pt} \ln P_{pt} + \beta_d \ln P_d + \beta_n \ln P_n + \beta_s \ln P_s + e \quad (1)$$

Using the data in Exercise 6.22, use IML and SAS to answer the following.

1. Estimate the model using least squares. Compute and report coefficient estimates, their standard errors, t-ratios, and their p-values.
2. Compute  $R^2$ .
3. Perform the overall F-test for regression significance.
4. Report your estimate of the scale parameter,  $\sigma^2$ .

## Problem 2

Using the same model and data do the following:

1. Test the following hypotheses at the 5% level using a Wald test (i.e., equation (6.4.7)):

- (a)  $\beta_{pt} = 0$  and  $\beta_Y = 1$ . Own price elasticity of demand is one and the income elasticity is zero.
  - (b)  $\beta_{nc} = \beta_{uc}$  and  $\beta_Y = 1$ .
  - (c)  $\beta_{nc} = \beta_{uc}$ ,  $\beta_n = \beta_d$ , and  $\beta_Y = 1$ .
2. Find the p-values associated with the Wald tests.
  3. Find the 3 sets of restricted estimates using the RLS estimator

$$b^* = b + (X'X)^{-1}R'[R(X'X)^{-1}R']^{-1}(r - Rb) \quad (2)$$

- as discussed in class. Compute and report standard errors, t-ratios, and p-values.
4. Repeat the tests using the other forms of the statistics  $\lambda_2$  and  $\lambda_3$  discussed in class.
  5. Find the restricted estimates and repeat the hypothesis tests using the RESTRICT and TEST statements in SAS's PROC AUTOREG.

Important Note: Whenever you do a hypothesis test, be sure to state the null and alternative hypotheses, the distribution of your test statistic under the null hypothesis, your decision criterion, and your decision.

### Problem 3

Write an IML module that computes the RESET test. Perform the RESET with powers 2 and 3 on the gasoline demand model above. Based on the outcome of this test, is your model misspecified? If so, reestimate the model in level form and repeat the RESET test.

Remember: Whenever you do a hypothesis test, be sure to state the null and alternative hypotheses, the distribution of your test statistic under the null hypothesis, your decision criterion, and your decision.

### Problem 4

Data on U.S. consumption and GNP for the years 1962-1982 appear in Table 7.8 of Greene. Consider the following model:

$$C_t = \alpha + \beta Y_t + \gamma C_{t-1} + e_t \quad (3)$$

In this model the short-run MPC is  $\beta$  and the long-run MPC is  $\delta = \beta/(1-\gamma)$ .

1. Test the null hypothesis that  $\delta = 1$  against the alternative that it is not at the 5% level.
2. Test the (joint) null hypothesis that  $\delta = 1$  and  $\beta = .5$  against the alternative  $\delta \neq 1$  or  $\beta \neq .5$  at the 5% level

## Problem 5

The Longley (1967) data is widely used to calibrate regression programs. The dependent variable in the model is employment and the regressors include a time trend (Year), the GNP deflator, GNP, size of the Armed forces, and a constant. The model appears below:

$$\text{Employment}_t = \beta_1 + \text{Year}_t\beta_y + \text{Deflator}_t\beta_d + \text{GNP}_t\beta_g + e_t \quad (4)$$

where  $t = 1, \dots, T$ . Investigate the parameter stability of the model using the CUSUM test and the Longley data found in Table 7.4 in Greene (1997).

1. Reproduce the results in Table 7.5 of Greene. Essentially, you will regress total employment on a constant, year, deflator, and GNP using the following different sample periods: 1947-1962, 1947-1953, and 1954-1962.
2. Perform the Chow test for parameter stability at the 5% level under the assumption that the subset equation variances are the same.
3. Perform the Chow test for parameter stability at the 5% level under the assumption that the subset equation variances are not the same. See equation 7-23 (Greene 1997).
4. Reproduce the results in (Greene 1997) Table 7.6. Hint: You will have to rearrange your data so that  $X$  and  $Y$  in the pooled model have 32 observations.
5. Compute the CUSUM and its upper and lower bounds using IML. You may confirm your results using PROC AUTOREG which permits you to output these statistics into a data set which can in turn be printed. There is an example of the SAS statements needed to accomplish this posted at the end of these exercises.

## Problem 6

Gaver and Geisel (1974) propose two forms of a consumption function:

$$H_0: C_t = \alpha + \beta Y_t + \gamma C_{t-1} + e_t \quad (5)$$

and

$$H_a: C_t = \alpha + \beta Y_t + \gamma Y_{t-1} + e_t \quad (6)$$

The first model states that consumption responds to changes in income over many periods (infinite distributed lag model) and the second states that it responds to changes in income over only two periods.

1. Using the data from the above problem, determine which of the above hypotheses are supported by the data using the J test and the Cox test. (Set the level of the test at 5%).
2. Compute the AIC and SC statistics for each model.

## Problem 7

Estimate the model in equation (3) using least squares and compute standard errors, t-ratios, and p-values using White's heteroscedasticity covariance matrix.

## Problem 8

1. Test for the presence of heteroscedasticity in model (3) using White's test.
2. Test for its presence again using the Breusch-Pagan test assuming that heteroscedasticity is a function of  $z_{t2} = C_t/GNP_t$  and a constant. First perform the test under the assumption of normality, then repeat the test under the assumption that the errors are not normally distributed.

## Problem 9

Use PROC AUTOREG to determine whether the model in equation (3) is autocorrelated of order 1.

## Programming Examples

```
DATA ONE;                                * creates a data set called ONE ;
INFILE 'c:\ex6_22.dat' firstobs=2;      * reads data, skip line 1;
input Y x2 x3 x4;                        * INPUT ;
    lx2=log(x2);                          * take natural logs of variables ;
    lx3=log(x3);
    lx4=log(x4);
    x4_lag=lag(x4);
    ly=log(y);
    int=1;

Proc AutoReg;
    model y = x2 x3 x4/reset;
    Test x2=x4,x3=0;
    output out=two recres=recur cusum=c1 cusumlb=lb cusumub=ub;
proc print data=two;

Proc Autoreg;
    model y = x2 x3 x4;
    restrict x2=x4,x3=0;

Proc IML;
    vname= { const, GasPrice, Income, NewCar};
    use one;
    read all var{y} into y;
    read all var{ int x2 x3 x4 } into x;
    read all var{ int x2 x3 x4 } into x where(x4 > 1963);

    t=nrow(x);
    k=ncol(x);
    b=inv(x'*x)*x'*y;
```

## References

- Gaver, K. and M. Geisel**, "Discriminating Among Alternative Models: Bayesian and Non-Bayesian Methods," *Frontiers in Econometrics*, 1974.
- Greene, William H.**, *Econometric Analysis*, 3rd ed., Prentice-Hall, 1997.
- Longley, J.**, "An Appraisal of Least Squares Programs from the Point of the User," *The Journal of the American Statistical Association*, 1967, 62, 819-841.