Econ 301-1 **Bilkent University**
Econometrics **Department of Economics**

Lab Exercise

The American households tend to travel when they take a vacation. The distance they travel for vacationing is modelled as follows:

$$miles_i = \beta_1 + \beta_2 income_i + \beta_3 age_i + \beta_4 kids_i + \varepsilon_i$$

where the distance of vacationing is a function of the family income, age of the head of the household, and the number of kids in the family.

a) What are the expected signs for the above coefficients?
b) Estimate the model.
c) Examine the residuals of the above estimation. Try plotting the residuals against age and income. (Choose X-Y graph and scatter diagram) What do you observe? If the same graphs are calculated for squares of the residuals do the results change? What can you say about the form of heteroscedasticity according to these plots?
d) Perform the Park's test. To do so
   Estimate $Log(\hat{\varepsilon}_i^2) = \alpha_0 + \alpha_1 \log(income_i) + u_i$, the test
   H_0: α_1=0 versus H_a: not H_0.
e) Perform a set of Glejser tests. To do so
   Estimate i. $|\hat{\varepsilon}_i| = \alpha_0 + \alpha_1 income_i + u_i$
   ii. $|\hat{\varepsilon}_i| = \alpha_0 + \alpha_1 \sqrt{income_i} + u_i$
   iii. $|\hat{\varepsilon}_i| = \alpha_0 + \alpha_1 (income_i)^2 + u_i$
   then for each cases test H_0: α_1=0 versus H_a: not H_0.
f) If you conclude that the assumption

$$var(\varepsilon_i) = \sigma^2 income_i$$

is valid in this data set, conduct a Goldfeld-Quant test according to the following formula:

$$GQ = \frac{\sum \hat{\varepsilon}_i^2 / (N_2 - K_2)}{\sum \hat{\varepsilon}_i^2 / (N_1 - K_1)}$$

Which will have a F-distribution of ($N_2 - K_2, N_1 - K_1$) degrees of freedom if the null hypothesis $H_0 : \sigma_1^2 = \sigma_2^2$ is correct. (do not forget to sort the data according to income)
(If GQ is >the F-critical then reject the null hypothesis that the error variances are equal across sub-samples)

g) Perform the White ***<u>test</u>***.
   Estimate $(\hat{\varepsilon}_i^2) = \alpha_0 + \alpha_1 income + \alpha_2 age + \alpha_3 kids + \alpha_4 income^2 + \alpha_5 age^2 + \alpha_6 kids^2 + \alpha_7 income * age + \alpha_8 income * kids + \alpha_9 kids * age + u_i$,
   the test Ho:$\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = \alpha_5 = \alpha_6 = \alpha_7 = \alpha_8 = \alpha_9 = 0$ versus Ha: Not Ho.
   Estimate the model get nR^2 compare with the chi-square
h) Perform Koenker-Basett Test.
   Estimate $(\hat{\varepsilon}_i^2) = \alpha_0 + \alpha_1 \widehat{miles} + u_i$, the test
   then test H_0: α_1=0 versus H_a: not H_0 by using the t-test.

i) Find the generalized (weighted) least square estimates of the coefficients if

i. $var(\varepsilon_i) = \sigma^2 income_i$
ii. $var(\varepsilon_i) = \sigma^2 income_i{}^2$
iii. $var(\varepsilon_i) = \sigma^2 / income_i$
iv. $var(\varepsilon_i) = \sigma^2 (income_i + kids_i^2)$
v. $var(\varepsilon_i) = \sigma^2 (income_i / age_i)$

j) Estimate the model with OLS but use White's ***robust standard errors*** for inferences.