44-6873-00L- **Econometrics for Business and Economics**	Dr Amr Algarhi (Miro) Department of Management Sheffield Hallam University
Exercise sheet **7. Dummy variables and multicollinearity** Week 27	

Question 1 (Stata).

Download the data set "***utown.dta***" file, which includes data to explain house prices. The variable $price$ is the house price (in thousands of dollars). The variable $sqft$ is the square feet of living area (in hundreds). The variable age is the house age in years.

The data set also includes the following dummy (indicator) variables that indicate the presence or absence of particular characteristics:

utown	= 1 if the house is close to a university (university town)
pool	= 1 if the house has a pool
fplace	= 1 if the house has a fireplace

(a) Estimate the following model which includes dummy variables in the explanatory variables.

$$price = \beta_1 + \beta_2 sqft + \beta_3 age + \delta_1 utown + \delta_2 pool + \delta_3 fplace + \gamma(sqft \times utown) + u$$

Hint:

```
reg price age i.pool i.fplace i.utown##c.sqft
```

(b) Test the significance of the University Town location.

Hint:

```
test 1.utown 1.utown#c.sqft
```

(c) Write down the estimated regression function for the houses near the university.

(d) Create a dummy variable to indicate large houses, more than 2500 square feet in size.

Hint:

```
gen large = (sqft>25)
```

Question 2 (Stata).

This question considers the problem of collinearity which makes it difficult or impossible to compute the parameter estimates and various other statistics with much precision. Collinearity arises because of poor experimental design or because of data.

Download the cars data set, after clearing any pervious data out of memory. The file, "***cars.dta***", contains the following four variables:

mpg	miles per gallon
cyl	number of cylinders
eng	engine displacement in cubic inches
wgt	vehicle weight in pounds

(a) Are the variables in this data set correlated?
Hint:
```
corr
```

(b) Regress the miles per gallon (mpg) on the explanatory variables (cyl, eng and wgt).

(c) Now, test a series of hypotheses. The first is for the significance of cyl, the second for the significance of eng, and the third is of their joint significance.
Hint:
```
test cyl
test eng
test cyl eng
```

(END)