EE 325 HW 4 Answer

1. From the data for 46 states in the United States for 1992, Baltagi obtained the following regression results

$$\widehat{\log C} = 4.30 - 1.34 \log P + 0.17 \log Y$$
$$se = (0.91) \quad (0.32) \qquad (0.20)$$
$$\bar{R}^2 = 0.27$$

Where C = cigarette consumption, packs per year
P = real price per pack
Y = real disposable income per capita
and the figures in the parentheses are the estimated standard errors

a. What is the elasticity of demand for cigarettes with respect to price? Is it statistically significant? If so, is it statistically different from 1?

The elasticity is –1.34. It is significantly different from zero, for the *t* value under the null hypothesis that the true elasticity coefficient is zero is:

$$t = \frac{-1.34}{0.32} = -4.4687$$

The *p value* of obtaining such a *t* value is extremely low.
However, the elasticity coefficient is different from one because under the null hypothesis that the true elasticity is 1, the t value is

$$t = \frac{-1.34 - 1}{0.32} = -7.3125$$

This *t* value is statistically significant.

b. What is the income elasticity of demand for cigarettes? Is it statistically significant?

The income elasticity is statistically different from zero.

2. From the sample of 209 firms, Wooldridge obtained the following regression results:

$$\widehat{\log(salary)} = 4.32 + 0.280 \log(sales) + 0.0174 roe + 0.00024 ros$$
$$se = (0.32) \qquad (0.035) \qquad (0.0041) \qquad (0.00054)$$
$$R^2 = 0.283$$

Where salary = salary of CEO
sales = annual firm sales
roe = return on equity in percent
ros = return on firm's stock

and the figures in the parentheses are the estimated standard errors

a. Interpret the preceding regression taking into account any prior expectations that you may have about the sings of the various coefficients.

A priori, salary and each of the explanatory variables are expected to be positively related, which they are. The partial coefficient of 0.280 means, *ceteris paribus*, the elasticity of CEO salary is a 0.28 percent. The coefficient 0.0174 means, *ceteris paribus*, if the rate of return on equity goes up by 1 percentage point (Note: not by 1 percent), then the CEO salary goes up by about 1.07 %. Similarly, *ceteris paribus*, if return on the firm's stock goes up by 1 percentage point, the CEO salary goes up by about 0.024%.

b. Which of the coefficients are individually statistically significant at the 5 percent level?

Under the *individual*, or separate, null hypothesis that each true population coefficient is zero, you can obtain the *t* values by simply dividing each estimated coefficient by its standard error. These *t* values for the four coefficients shown in the model are, respectively, 13.5, 8, 4.25, and 0.44. Since the sample is large enough, using the two-*t* rule of thumb, you can see that the first three coefficients are individually highly statistically significant, whereas the last one is insignificant.

c. What is the overall significance of the regression? Which test do you use?
To test the overall significance, that is, all the slopes are equal to zero, use the *F* test

$$F = \frac{R^2/(k-1)}{(1-R^2)/(n-k)} = \frac{0.283/3}{(0.717)/205} = 27.02$$

Under the null hypothesis, this *F* has the *F* distribution with 3 and 205 df in the numerator and denominator, respectively. The *p value* of obtaining such an *F* value is extremely small, leading to rejection of the null hypothesis.

3. Data on advertising impressions retained and advertising expenditure for a sample of 21 firms. Letting Y represent impressions retained and X the advertising expenditure, the following regressions were obtained:

$$\begin{aligned}
&\textit{Model I} && \hat{Y}_i = 22.163 + 0.3631X_i \\
& && se = (7.089) \quad (0.0971) \\
& && r^2 = 0.424 \\
&\textit{Model II} && \hat{Y}_i = 7.059 + 1.0847X_i - 0.0040X_i^2 \\
& && se = (9.986) \quad (0.3699) \quad (0.0019) \\
& && R^2 = 0.53
\end{aligned}$$

a. Interpret both models

In Model I the slope coefficient tells us that per unit increase in the advertising expenditure, on average, retained impressions go up by 0.363 units. In Model II the (average) rate of increase in retained impressions depend on the level of advertising expenditure. Taking the derivative of Y with respect to X, you will obtain:

$$\frac{dY}{dX} = 1.0847 - 0.008X$$

This would suggest that retained impressions increase at a decreasing rate as advertising expenditure increases.

b. Which is a better model? Why?

c. Which statistical test(s) would you use to choose between the two models?

We can treat Mode I 1 as the abridged, or restricted, version of Model II and hence can use the restricted least-squares technique to decide between the two models. Since the dependent variable in the two models is the same, we can use the *R*2 version of the *F* test given in (8.7.10). The results are as follows:

$$F = \frac{(0.53 - 0.424)/1}{(1 - 0.53)/18} = \frac{0.106}{0.0261} = 4.0613$$

Under the usual assumptions of the *F* test, the preceding *F* value follows the *F* distribution with 1 df and 18 df in the numerator and denominator, respectively. For these dfs the critical *F* value is 4.41 (5% level) and 3.01 (10% level).; the *p value* is 0.0591 or about 6%, which is close to 5%. It seems that we should retain the squared X variable in the model.

d. Are there "diminishing returns" to advertising expenditure, that is, after a certain level of advertising expenditure (the saturation level), does it not pay to advertise? Can you find out what the level of expenditure might be?

As noted in (*b*), there are diminishing returns to advertising expenditure; if the coefficient of the *X*-squared term were positive, there would have been increasing returns to advertising. Equating the derivative in (*b*) to zero, we obtain: 1.0847=0.0008X, which gives X =135.58. At this value of X the rate of increase of Y with respect to X is zero. Since *X* is measured in millions of dollars, we can say that at the level of expenditure of about 136 millions of dollars there is no further gain in retained impressions, which are measured in millions of impressions.

Empirical exercise using STATA ☺☺

4. *Disposable income and personal savings* The Chow test was introduced to see if a structural change occurred within the data between the two time periods. Table 8.11 includes updated data containing the values from 1970-2005. According to the National Bureau of Economic Research, the most recent U.S. business contraction cycle ended in late 2001.

   a. Estimate both the model for the full dataset (years 1970-2005) and 1970-1981

Time period: 1970-2005

Source	SS	df	MS
Model	3071.63945	1	3071.63945
Residual	265277.864	34	7802.29012
Total	268349.503	35	7667.12867

Number of obs	=	36
F(1, 34)	=	0.39
Prob > F	=	0.5346
R-squared	=	0.0114
Adj R-squared	=	-0.0176
Root MSE	=	88.331

savings	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
income	.0037239	.005935	0.63	0.535	-.0083375	.0157852
_cons	184.8291	27.89717	6.63	0.000	128.1352	241.5229

Time period: 1970-1981

Source	SS	df	MS
Model	27553.7951	1	27553.7951
Residual	1613.75489	10	161.375489
Total	29167.55	11	2651.59545

Number of obs	=	12
F(1, 10)	=	170.74
Prob > F	=	0.0000
R-squared	=	0.9447
Adj R-squared	=	0.9391
Root MSE	=	12.703

savings	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
income	.1017888	.0077898	13.07	0.000	.084432	.1191456
_cons	-5.707531	11.03869	-0.52	0.616	-30.30326	18.8882

Time period: 1982-2005

Source	SS	df	MS
Model	83030.0418	1	83030.0418
Residual	69568.4388	22	3162.20176
Total	152598.481	23	6634.71655

Number of obs	=	24
F(1, 22)	=	26.26
Prob > F	=	0.0000
R-squared	=	0.5441
Adj R-squared	=	0.5234
Root MSE	=	56.233

savings	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
income	-.0300647	.0058672	-5.12	0.000	-.0422326	-.0178968
_cons	394.3411	33.26142	11.86	0.000	325.3611	463.3211

   b. Using Chow test, determine if there is a significant difference between 1970-1981 and the full dataset.

The Chow test for this question is constructed as follows:

$$RSS_{UR} = RSS_{pre1981} + RSS_{post1981} = 71182.195$$

$$F = \frac{(RSS_R - RSS_{UR})/k}{(RSS_{UR})/(n_1 + n_2 + k)} = \frac{(265277.871 - 71182.195)/2}{(71182.195)/(32)} = 46.3547$$

With 2 and 32 degrees of freedom, the *F* table in the book gives a 1% critical value of 5.39 (using denominator df of 30, which is more conservative). Therefore, this supports the idea that there was a structural break between the two time periods.