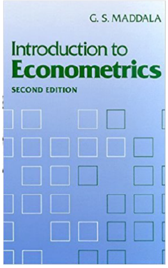

G. S. Maddala, *Introduction to econometrics*, 1992

8.10 Illustrative Example

Consider the data in Table 8.4. The data refer to a study on the deterrent effect of capital punishment by McManus.[29] The data are cross-sectional data for 44 states in the United States in 1950. There are two dummy variables in the data. D_2, which is a South–North dummy, is clearly an explanatory variable. But D_1 can be both an explained and explanatory variable. If it is an explained variable we would consider it as "a propensity to have capital punishment." ·

Let us first consider the regression of M on all the other variables. The results are as follows: (Figures in parentheses are *t-ratios*, not standard errors.)

$$\begin{aligned}\hat{M} = & \underset{(-0.82)}{-8.50} - \underset{(-1.38)}{3.696PC} - \underset{(-0.54)}{3.568PX} + \underset{(2.11)}{2.598D_1} \\ & \underset{(2.62)}{-0.018T} - \underset{(-2.3)}{4.095Y} + \underset{(1.82)}{0.400LF} \\ & + \underset{(1.17)}{6.444NW} + \underset{(1.93)}{2.541D_2} \qquad R^2 = 0.7746\end{aligned}$$

Some of the coefficients have signs opposite to those we would expect.

Let us now consider treating D_1 as an explained variable. We will consider T, Y, LF, NW, and D_2 as the explanatory variables.

The linear probability model gave the following results (figures in parentheses are t-ratios obtained from an ordinary regression program that ignores the zero-1 characteristic of the dependent variable):

$$\begin{aligned}\hat{D}_1 = & \underset{(1.50)}{1.993} + \underset{(1.46)}{0.00146T} + \underset{(2.74)}{0.658Y} - \underset{(-1.93)}{0.055LF} \\ & + \underset{(2.62)}{1.988NW} + \underset{(1.81)}{0.343D_2} \qquad R^2 = 0.3376\end{aligned}$$

What these results indicate is that southern states and states with higher percentages of nonwhites have a positive effect on the probability of having capital punishment. The percentage of labor force employed has a negative effect on the probability of having capital punishment. What is perplexing is the coefficient of Y (median family income), which is significantly positive. One possible explanation for this is that states with high incomes (New York, California, etc.) also have big cities where crime rates are high.

[29] Walter S. McManus, "Estimates of the Deterrent Effect of Capital Punishment: The Importance of the Researcher's Prior Beliefs," *Journal of Political Economy*, Vol. 93, April 1985, pp. 417–425.

Let us now look at the logit and probit estimates.[30] The logit model gave (figures in parentheses are asymptotic t-ratios)

$$D_1 = \underset{(0.53)}{10.99} + \underset{(1.87)}{0.0194T} + \underset{(1.88)}{10.61Y} - \underset{(-1.40)}{0.668LF} + \underset{(1.95)}{70.99NW} + \underset{(0.02)}{13.33D_2}$$

The results of the probit model were (figures in parentheses are asymptotic t-ratios)

$$\hat{D}_1 = \underset{(0.61)}{6.92} + \underset{(2.00)}{0.0113T} + \underset{(2.05)}{6.46Y} - \underset{(-1.59)}{0.409LF} + \underset{(2.05)}{42.50NW} + \underset{(0.04)}{4.63D_2}$$

As mentioned earlier, the logit coefficients have to be divided by 1.6 to be comparable to the probit coefficients. Such division produces the coefficients 6.87, 0.0121, 6.63, −0.418, 44.37, and 8.33, respectively which are close to the probit coefficients. Surprisingly, D_2 is not significant, but all the other coefficients have the same signs as in the linear probability model. The coefficient of Y is still positive and is significant.

Table 8.4 Determinants of Murder Rates in the United States (Cross-Section Data on States in 1950)[a]

N	M	PC	PX	D_1	T	Y	LF	NW	D_2
1	19.25	0.204	0.035	1	47	1.10	51.2	0.321	1
2	7.53	0.327	0.081	1	58	0.92	48.5	0.224	1
3	5.66	0.401	0.012	1	82	1.72	50.8	0.127	0
4	3.21	0.318	0.070	1	100	2.18	54.4	0.063	0
5	2.80	0.350	0.062	1	222	1.75	52.4	0.021	0
6	1.41	0.283	0.100	1	164	2.26	56.7	0.027	0
7	6.18	0.204	0.050	1	161	2.07	54.6	0.139	1
8	12.15	0.232	0.054	1	70	1.43	52.7	0.218	1
9	1.34	0.199	0.086	1	219	1.92	52.3	0.008	0
10	3.71	0.138	0	0	81	1.82	53.0	0.012	0
11	5.35	0.142	0.018	1	209	2.34	55.4	0.076	0
12	4.72	0.118	0.045	1	182	2.12	53.5	0.299	0
13	3.81	0.207	0.040	1	185	1.81	51.6	0.040	0
14	10.44	0.189	0.045	1	104	1.35	48.5	0.069	1
15	9.58	0.124	0.125	1	126	1.26	49.3	0.330	1
16	1.02	0.210	0.060	1	192	2.07	53.9	0.017	0
17	7.52	0.227	0.055	1	95	2.04	55.7	0.166	1
18	1.31	0.167	0	0	245	1.55	51.2	0.003	0
19	1.67	0.120	0	0	97	1.89	54.0	0.010	0
20	7.07	0.139	0.041	1	177	1.68	52.2	0.076	0
21	11.79	0.272	0.063	1	125	0.76	51.1	0.454	1
22	2.71	0.125	0	0	56	1.96	54.0	0.032	0
23	13.21	0.235	0.086	1	85	1.29	55.0	0.266	1
24	3.48	0.108	0.040	1	199	1.81	52.9	0.018	0
25	0.81	0.672	0	0	298	1.72	53.7	0.038	0

[30]The logit and probit estimates were computed using William Greene's LIMDEP program.

Table 8.4 (*Cont.*)

N	M	PC	PX	D_1	T	Y	LF	NW	D_2
26	2.32	0.357	0.030	1	145	2.39	55.8	0.067	0
27	3.47	0.592	0.029	1	78	1.68	50.4	0.075	0
28	8.31	0.225	0.400	1	144	2.29	58.8	0.064	0
29	1.57	0.267	0.126	1	178	2.34	54.5	0.065	0
30	4.13	0.164	0.122	1	146	2.21	53.5	0.065	0
31	3.84	0.128	0.091	1	132	1.42	48.8	0.090	1
32	1.83	0.287	0.075	1	98	1.97	54.5	0.016	0
33	3.54	0.210	0.069	1	120	2.12	52.1	0.061	0
34	1.11	0.342	0	0	148	1.90	56.0	0.019	0
35	8.90	0.133	0.216	1	123	1.15	56.2	0.389	1
36	1.27	0.241	0.100	1	282	1.70	53.3	0.037	0
37	15.26	0.167	0.038	1	79	1.24	50.9	0.161	1
38	11.15	0.252	0.040	1	34	1.55	53.2	0.127	1
39	1.74	0.418	0	0	104	2.04	51.7	0.017	0
40	11.98	0.282	0.032	1	91	1.59	54.3	0.222	1
41	3.04	0.194	0.086	1	199	2.07	53.7	0.026	0
42	0.85	0.378	0	0	101	2.00	54.7	0.012	0
43	2.83	0.757	0.033	1	109	1.84	47.0	0.057	1
44	2.89	0.356	0	0	117	2.04	56.9	0.022	0

[a] N, observation number; M, murder rate per 100,000, FBI estimate 1950; PC, (number of convictions/number of murders) in 1950; PX, average number of executions during 1946–1950 divided by convictions in 1950; Y, median family income of families in 1949 (thousands of dollars); LF, labor force participation rate 1950 (expressed as a percent); NW, proportion of population that is nonwhite in 1950; D_2, dummy variable, 1 for southern states, 0 for others; D_1, dummy variable which is 1 if the state has capital punishment, 0 otherwise ($D_1 = 1$ if $PX > 0$, 0 otherwise); T, median time served in months of convicted murderers released in 1951.

One other problem is that, as mentioned in Section 8.9, the coefficients of the logit model should be approximately four times the coefficients of the linear probability model, but the coefficients we have obtained are much higher than that. One possible reason for this is the poor fit given by the linear probability model. To investigate this we computed the different measures of R^2's discussed in the preceding section, and the R^2's for the linear probability model are significantly lower than those for the logit and probit models.

In Table 8.5 we present four different measures of R^2's.[31] The first two are easy to compute and are reasonable measures of R^2's. The measures suggested by Cragg and Uhler and by McFadden both depend on the computation of L_R and L_{UR}. The results indicate that there is not much to choose between the logit

[31] We did not compute Amemiya's R^2's. Although he has given an expression for the residual sum of squares, he has not given an expression for the total sum of squares (which should also be appropriately weighted). Using the unweighted total sum of squares $\Sigma(y_i - \bar{y})^2$ produces a negative R^2.

Table 8.5 Different R^2 Measures for the Logit, Probit, and Linear Probability Models

	Logit	*Probit*	*Linear Probability*
Squared correlation between D_1 and $\hat{D}_1$	0.6117	0.6099	0.3376
Effron's R^2	0.6116	0.6095	0.3376
Cragg–Uhler's R^2	0.7223	0.7258	0.5273
McFadden's R^2	0.6083	0.6124	0.4029

and probit models and that both are better than the linear probability model. From the practical point of view it appears that the squared correlation between D_1 and $\hat{D}_1$ and Effron's R^2 are sufficient for many problems.

Since we decided on the probit and logit models and D_2 was not significant in these models, we decided to drop that variable and reestimate the probit and logit models. The revised estimates were (figures in parentheses are asymptotic t-ratios)

Logit

$$\hat{D}_1 = \underset{(0.84)}{16.57} + \underset{(1.72)}{0.0165T} + \underset{(1.81)}{9.13Y} - \underset{(-1.49)}{0.715LF} + \underset{(2.38)}{85.36NW}$$

$$R^2(D_1, \hat{D}_1) = 0.5982 \qquad \text{Effron's } R^2 = 0.5982$$
$$\text{Cragg–Uhler's } R^2 = 0.7077 \qquad \text{McFadden's } R^2 = 0.5914$$

Probit

$$\hat{D}_1 = \underset{(0.98)}{10.27} + \underset{(1.86)}{0.0094T} + \underset{(1.97)}{5.55Y} - \underset{(-1.7)}{0.437LF} + \underset{(2.50)}{50.25NW}$$

$$R^2(D_1, \hat{D}_1) = 0.5950 \qquad \text{Effron's } R^2 = 0.5947$$
$$\text{Cragg–Uhler's } R^2 = 0.7113 \qquad \text{McFadden's } R^2 = 0.5955$$

Again, to make the logit coefficients comparable to the probit coefficients, we have to divide the former by 1.6. This gives 10.36, 0.0103, 5.71, -0.447, and 53.35, respectively, which are close to the probit coefficients.