Econometrics 1

Lecture 19

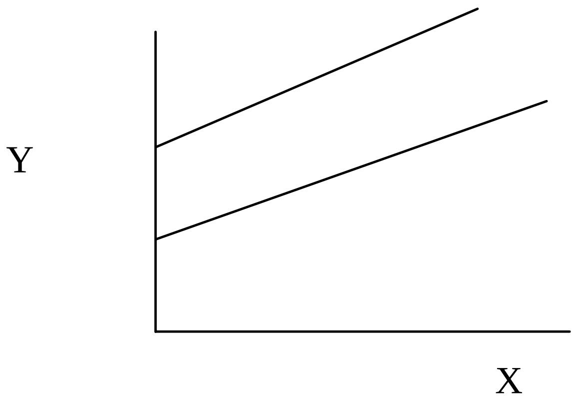
Dummy Dependent Variables Models

Dummy Variables

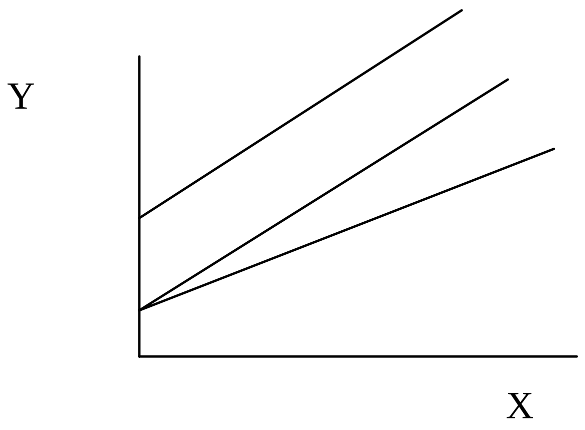
- We use *dummy variables*, which are explanatory variables that only take two values, usually 0 and 1.
- These simple variables are very powerful tools for capturing qualitative characteristics of individuals, such as gender, race, geographic region of residence.
- In general, we use dummy variables to describe any event that has only two possible outcomes.

There are mainly three roles for dummy variables as a reggressor

- Intercept dummy
- Slope dummy
- Slope and intercept interaction dummy



Intercept dummy



Intercept and slope dummy

Use of Dummy in House Price Model

Consider that the price of house depends on its size

$$P_t = \beta_1 + \beta_2 S_t + e_t$$

It is common experience that price of houses depend on its location.

A dummy variable *D* can be used to pick up the location effec

$$D_t = \begin{cases} 1 \text{ if property is in the desirable neighbourhood} \\ 0 \quad \text{if property is not in the desirable area} \end{cases} \qquad (2$$

- Adding this dummy variable to the regression model, along with a new parameter δ, we can write (1) as

$$P_t = \beta_1 + \delta D_t + \beta_2 S_t + e_t$$

Impact of Location in House Price

The expected price now is higher in the urban areas than in non-urban locations

$$E(P_t) = \begin{cases} (\beta_1 + \delta) + \beta_2 S_t + e_t & if \quad D_t = 1 \\ \beta_1 + \beta_2 S_t + e_t & if \quad D_t = 0 \end{cases}$$

The expected price now is higher in the urban areas than in non-urban locations

$$E(P_t) = \begin{cases} (\beta_1 + \delta) + \beta_2 S_t + e_t & if \quad D_t = 1 \\ \beta_1 + \beta_2 S_t + e_t & if \quad D_t = 0 \end{cases}$$

(3)

Adding the dummy variable D_t to the regression model creates a *parallel shift* in the relationship by the amount δ. This is an example of a intercept dummy.

Slope Dummy

It explains interaction between a continuous variable and a dummy variable. The size of house has more influence on prices in the urban than in rural areas

$$P_t = \beta_1 + \beta_2 S_t + \gamma \left(S_t D_t \right) + e_t$$

(4)

The effect of a change in house size on price is.

$$\frac{\partial E\left(P_t\right)}{\partial S_t} = \begin{cases} \beta_2 + \gamma & if \quad D_t = 1 \\ \beta_2 & if \quad D_t = 0 \end{cases}$$

(5)

The slope and the intercept dummy: starting price may be higher in the urban areas and may grow at a higher rate

$$P_t = \beta_1 + \delta D_t + \beta_2 S_t + \gamma \left(S_t D_t \right) + e_t$$

$$E\left(P_t\right) = \begin{cases} \left(\beta_1 + \delta\right) + \left(\beta_2 + \gamma\right) S_t & if \quad D_t = 1 \\ \beta_1 + \beta_2 S_t & if \quad D_t = 0 \end{cases}$$

Applications and Dummy Variable Trap

- ***Interactions Between Qualitative Factors***

- ***Seasonal Dummies***
- ***Annual Dummies***
- ***Regime Effects and structural breaks***
- ***BHPS is a rich source to explore various relations with dummy variable (for access to BHPS check)***
- Dummy variable trap

Chow Test

Test for stability of parameters or structural change (Chow test)

Use n_1 and n_2 observations to estimate overall and separate

regressions with (n_1+n_2-k, n_1-k, and n_2-k) dfs; obtain SSR_1(with n_1+n_2-k dfs),

SSR_2(with n_1-k dfs), SSR_3 (with n_2-k dfs) and SSR_4 = SSR_1+ SSR_2 (with

n_1+n_2-2k dfs), obtain $S_5 = S_1 - S_4$; do $F = \frac{S_5 / k}{S_4 / (n_1 + n_1 - 2k)}$.

The advantage of this approach to the Chow test is that it does not require the construction of the dummy and interaction variables.

A Shazam Programme

```
sample 1 40
*be=expenditure on beer, s=1 male 0 =female,
e1=high school,
*e2=college e3=university y=income, age
read be s e1 e2 e3 y age
*generate the interaction variable for gender and
income and age and income
genr sy=s*y
genr ay = age*y
ols be s y sy/pcov cov=cov
*
* compute se's for female function
*
read r/rows=2 cols=4
 0 1 1 0
 1 0 0 1
matrix v=r*cov*r'
```

```
matrix se=sqrt(diag(v))
print v se
sample 1 21
ols be y
sample 22 40
ols be y
sample 1 40
genr h=1-s
genr hy=h*y
ols be h s hy sy/noco
ols be s e2 e3 y sy age
confid  s age /gnu
stop
```