Lecture notes VII

Non linear regression :
(Stock and Watson ch. 8)

general non-linear regression ~~function~~:

$$Y = f(X_1, \dots, X_k) + u$$

- $f(X_1, \dots, X_k)$ is regression function
  $E(Y | X_1, \dots, X_k) = f(X_1, \dots, X_k)$
- u is error term: $E(u | X_1, \dots, X_k) = 0$
  Note: u is still additive

The effect on Y of a partial change in X_1 in a non-linear model is

$$\Delta Y \equiv \underbrace{f(X_1 + \Delta X_1, X_2, \dots, X_k)}_{= E(Y | X_1 + \Delta X_1, X_2, \dots, X_k)} - \underbrace{f(X_1, \dots, X_k)}_{= E(Y | X_1, \dots, X_k)}$$

$$\simeq \frac{\partial f(X_1, \dots, X_k)}{\partial X_1} \cdot \Delta X_1$$

That is ΔY is the change in $E(Y | X_1, \dots, X_k)$ when X_1 changes from X_1 to $X_1 + \Delta X_1$

In the linear model:

$$f(X_1,\dots,X_k) = \beta_0+\beta_1 X_1+\dots+\beta_k X_k$$

and

$$\Delta Y = \underbrace{\frac{\partial f(X_1,\dots,X_k)}{\partial X_1}}_{=\beta_1}\cdot\Delta X_1 = \beta_1\cdot\Delta X_1$$

<u>Example 1</u>

(1) $$\ln Y = \underbrace{\beta_0+\beta_1 S+\beta_2 E+\beta_3 E^2}_{=f(S,E)} + u$$

where Y is earnings (of individual i)
S is years of schooling exceeding 7 years
E is years of experience

Non-linear relation between $\ln Y$ and E

$$\frac{\partial f(S,E)}{\partial E} = \beta_2+2\beta_3 E$$

, i.e. non-constant marginal effects

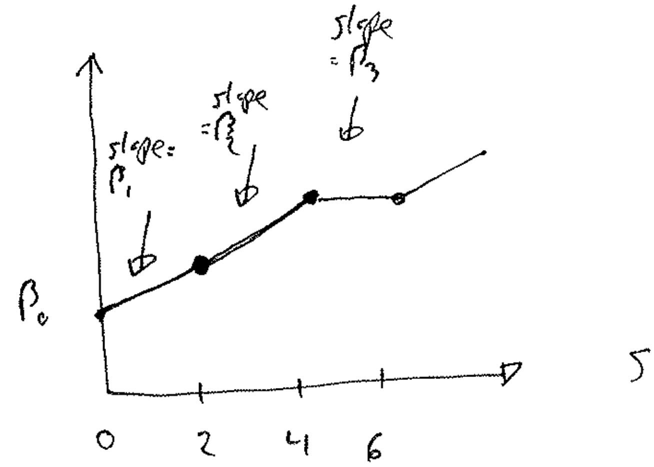
Note that (1) is <u>linear</u> in the unknown parameters! $\Rightarrow$ can be estimated by OLS

Example 2

Assume $f(S,E)$ is non-linear in schooling S. In particular:

$$\frac{\partial f(S,E)}{\partial S} = \begin{cases} \beta_1 & \text{if} \quad 0 \le S < 2 \\ \beta_2 & \text{if} \quad 2 \le S < 4 \\ \vdots & \\ \beta_6 & \text{is} \quad 10 \le S < 12 \end{cases}$$

$E(\ln Y \mid S, E=0)$



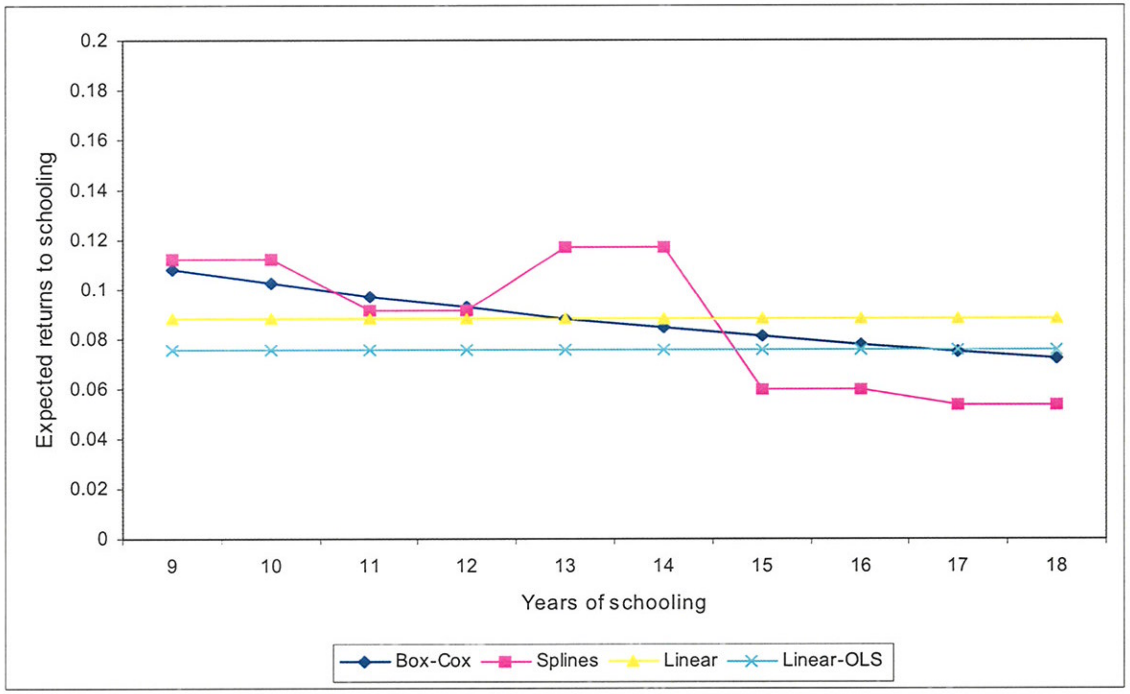
Regression model can be written as

$$\ln Y = \beta_0 + \beta_1 X_1 + \cdots + \beta_6 X_6 + \beta_7 E + \beta_8 E^2 + u$$

where

$$X_1 = \min(S,2)$$
$$X_2 = (\min(S,4) - 2) \times I(S \geq 2)$$
$$X_3 = (\min(S,6) - 4) \cdot I(S \geq 4)$$
$$\vdots$$
$$X_6 = (\min(S,12) - 10)\, I(S \geq 10$$

note:
$$I(S \geq k) = \begin{cases} 1 & \text{if } S \geq k \\ 0 & \text{else} \end{cases}$$



Expected marginal returns to schooling

The "average treatment effect" of schooling: increase in earnings of an additional year schooling

Here $g(S, E) = \beta_0 + \beta_1 X_1 + \dots + \beta_6 X_6 + \beta_7 E + \beta_8 E^2$

Can test whether model is linear in E by testing

a) $H_0: \beta_8 = 0$ vs $H_1: \beta_8 \neq 0$

And whether linear in S by the testing

b) $H_0: \beta_1 = \beta_2 = \dots = \beta_6$ vs H_1: "H_0 not true"

a) is a simple t-test

b) is an F-test:

$$F = \frac{(SSR_{restricted} - SSR_{unrestricted})/5}{SSR_{unrestricted}/(n-9)}$$

Note: 9 unrestricted parameters in

$$g(S, E) = \beta_0 + \beta_1 X_1 + \dots + \beta_6 X_6 + \beta_7 E + \beta_8 E^2$$

Note: Under H_0 $F \sim F_{5,\infty}$ when $n \to \infty$ ("large n")

The above examples are models that are non-linear in the <u>explanatory variables</u> (non-constant marginal effects) but linear in the <u>parameters</u>, and hence may be estimated by OLS

<u>Non-linear models which are non-linear in the parameters</u>:

Example: $Y_i = \begin{cases} 1 & \text{firm } i \text{ gets R\&D subsidy} \\ 0 & \text{else} \end{cases}$

If $Y = \beta_0 + \beta_1 X_1 + \dots + \beta_k X_k + u$

then $u = \begin{cases} 1 - \beta_0 - \beta_1 X_1 - \dots - \beta_k X_k & \text{if } Y = 1 \\ -\beta_0 - \beta_1 X_1 - \dots - \beta_k X_k & \text{if } Y = 0 \end{cases}$

cannot satisfy A.1 of OLS. Why?

Moreover: $E(Y|X_1, \dots, X_k)$ may not be between 0 and 1

Why is this problematic?

Alternative model:

$$E(Y|X_1, \dots, X_k) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 X_1 + \dots + \beta_k X_k)}} \in (0,1),$$

i.e. <u>logit</u>

Specific non-linear models

① $$Y_i = \beta_0 + \beta_1 X_i + \beta_2 X_i^2 + \dots + \beta_r X_i^r + \sum_{j=1}^{m} \lambda_j Z_{ji} + u_i$$

where $Z_{1i}, \dots, Z_{mi}$ are the other regressors and X_i enters non-linearly in a polynomial of order r

$r = 2$: quadratic model in X_i
$r = 3$: cubic model in X_i

① can also be estimated by OLS, since linear in the parameters

Testing the polynomial of degree r versus a linear alternative

H_0: $\beta_2 = 0$, $\beta_3 = 0$, $\vdots$, $\beta_r = 0$ vs. H_1: At least one of $\beta_2, \dots, \beta_r$ is different from zero

F-statistic $$\frac{(SSR_{restricted} - SSR_{unrestricted})/(r-1)}{SSR_{unrestricted}/(n-r-m-1)}$$

i) $F_{r-1,\infty}$ - distributed as $n \to \infty$ under H_0

How to find the degree of the polynomial ?

Algorithm:

1. Start wit a maximum order r. In practice $r = 2, 3, 4$ is useful.

2. Test $H_0: \beta_r = 0$ vs. $\beta_r \neq 0$

   If H_0 is rejected $\Rightarrow$ let r be the ~~order~~ degree of polynomial

   If H_0 not rejected, go on and test

   $H_0: \beta_{r-1} = 0$ vs. $\beta_{r-1} \neq 0$

   If H_0 is rejected $\Rightarrow$ let $r-1$ be the degree

   If H_0 is not rejected $\Rightarrow$ test

   $H_0: \beta_{r-2} = 0$ vs $\beta_{r-2} \neq 0$

   ~~etc~~.

   and so on....

Problem with polynomials: $X, X^2, \ldots, X^r$ tend to be highly collinear $\Rightarrow$ imprecise estimates

Logarithms

$\ln(e^{x}) = x \quad , e = 2.71828\cdots$

Changes on logarithmic scale can be interpreted as relative chance:

Let $y \equiv \ln Y$

$$\underbrace{\Delta y}_{\substack{= \text{change} \\ \text{on logarithmic} \\ \text{scale}}} = \ln(Y + \Delta Y) - \ln(Y) \simeq \underbrace{\frac{\Delta Y}{Y}}_{\substack{= \text{relative} \\ \text{change}}} \text{. Why?}$$

$\rightarrow 100 \cdot \frac{\Delta Y}{Y}$ is percentage change

3 logarithmic regression models

I Linear-log model:

$$Y = \beta_0 + \beta_1 \ln X + u$$

$$\Downarrow$$

$$\Delta Y = \underbrace{\frac{\beta_1}{X_1}}_{= \frac{\partial f(X_1)}{\partial X_1}} \cdot \Delta X = \beta_1 \frac{\Delta X_1}{X_1}$$

$$= \underbrace{\beta_1 \cdot 0.01}_{} \times \left(\frac{\Delta X_1}{X_1} \cdot 100 \right).$$

That is $\beta_1 \cdot 100 \equiv$ effect of 1 % change in X_1

II log-linear model:

$$\ln Y = \beta_0 + \beta_1 X + u$$

$$\Delta \ln Y = \beta_1 \, \Delta X \quad \underline{\text{or}}$$

$$\frac{\Delta Y}{Y} = \beta_1 \cdot \Delta X_1$$

If $\Delta X_1 = 1$, the relative change in Y is β_1

— or percentage change is $\underline{100 \cdot \beta_1}$

III log-log model:

$$\ln Y = \beta_0 + \beta_1 \ln X + u$$

$$\Downarrow$$

$$\Delta \ln Y = \frac{\partial f(x)}{\partial X} \cdot \Delta X = \frac{\beta_1}{X} \, \Delta X \quad \underline{\text{or}}$$

$$\frac{\Delta Y}{Y} = \beta_1 \cdot \frac{\Delta X}{X}$$

$$\Updownarrow$$

$$\underbrace{100 \cdot \frac{\Delta Y}{Y}}_{\text{percentage change in } Y} = \beta_1 \cdot \underbrace{100 \cdot \frac{\Delta X}{X}}_{\text{percentage change in } X}.$$

<u>That is:</u> β_1 is percentage change in Y when X changes with 1 percent — i.e. elasticity

Interaction terms

I Between dummy (binary) variables:

(1) $$Y = \beta_0 + \beta_1 D_1 + \beta_2 D_2 + u$$

$$D_1 = \begin{cases} 1 & \text{if father has university exam} \\ 0 & \text{else} \end{cases}$$

$$D_2 = \begin{cases} 1 & \text{if mother has university exam} \\ 0 & \text{else} \end{cases}$$

$$E(Y \mid D_1, D_2) = \beta_0 + \beta_1 D_1 + \beta_2 D_2$$

The marginal effect of D_1 is:

$$\frac{\partial E(Y \mid D_1, D_2)}{\partial D_1} = \beta_1$$

Include now the interaction term $D_1 \cdot D_2$ in (1):

$$Y = \beta_0 + \beta_1 D_1 + \beta_2 D_2 + \beta_3 \underbrace{D_1 D_2}_{\text{interaction}} + u$$

Then

$$\frac{\partial E(Y \mid D_1, D_2)}{\partial D_1} = \begin{cases} \beta_1 & \text{if } D_2 = 0 \\ \beta_1 + \beta_3 & \text{if } D_2 = 1 \end{cases}$$

Reasonable that $\beta_1 > 0$, $\beta_1 + \beta_3 > 0$ but $\beta_3 < 0$

II Interaction between dummy and continuous variable:

Example: Y is earnings, S schooling and

$$D = \begin{cases} 1 & \text{if woman} \\ 0 & \text{else} \end{cases}$$

$$\ln Y = \beta_0 + \beta_1 D + \beta_2 S + \beta_3 D \cdot S + u$$

$$E(\ln Y \mid D, S) = \beta_0 + \beta_1 D + \beta_2 S + \beta_3 DS$$

$$\frac{\partial E(\ln Y \mid D, S)}{\partial S} = \beta_2 + \beta_3 D = \begin{cases} \beta_2 & \text{if } D = 0 \\ \beta_2 + \beta_3 & \text{if } D = 1 \end{cases}$$

III Interaction between 2 continuous variables:

Example: $Y = \ln$ Output

$K =$ capital

$L =$ labour input (man hours)

$$\ln Y = \beta_0 + \beta_1 \ln K + \beta_2 \ln L + \beta_3 \ln K * \ln L + u$$

$$\frac{\partial E(\ln Y \mid K, L)}{\partial \ln L} = \beta_2 + \beta_3 \ln K$$

— the marginal product of labour depends on capital