

Το πρόβλημα των παραλειπόμενων μεταβλητών

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + u, \quad [3.40]$$

$$wage = \beta_0 + \beta_1 educ + \beta_2 abil + u. \quad [3.42]$$

$$\tilde{y} = \tilde{\beta}_0 + \tilde{\beta}_1 x_1. \quad [3.41]$$

$$wage = \beta_0 + \beta_1 educ + v,$$

$\tilde{\beta}_1 = \hat{\beta}_1 + \hat{\beta}_2 \tilde{\delta}_1$, where $\hat{\beta}_1$ and $\hat{\beta}_2$ are the slope estimators (if we could have them) from the multiple regression

$$y_i \text{ on } x_{i1}, x_{i2}, i = 1, \dots, n \quad [3.43]$$

and $\tilde{\delta}_1$ is the slope from the simple regression

$$x_{i2} \text{ on } x_{i1}, i = 1, \dots, n. \quad [3.44]$$

$$\begin{aligned} E(\tilde{\beta}_1) &= E(\hat{\beta}_1 + \hat{\beta}_2 \tilde{\delta}_1) = E(\hat{\beta}_1) + E(\hat{\beta}_2) \tilde{\delta}_1 \\ &= \beta_1 + \beta_2 \tilde{\delta}_1, \end{aligned} \tag{3.45}$$

which implies the bias in $\tilde{\beta}_1$ is

$$\text{Bias}(\tilde{\beta}_1) = E(\tilde{\beta}_1) - \beta_1 = \beta_2 \tilde{\delta}_1. \tag{3.46}$$

$$\text{plim } \tilde{\beta}_1 = \beta_1 + \beta_2 \delta_1, \tag{5.5}$$

where

$$\delta_1 = \text{Cov}(x_1, x_2) / \text{Var}(x_1). \tag{5.6}$$

TABLE 3.2 Summary of Bias in $\tilde{\beta}_1$ when x_2 Is Omitted in Estimating Equation (3.40)

	Corr(x_1, x_2) > 0	Corr(x_1, x_2) < 0
$\beta_2 > 0$	Positive bias	Negative bias
$\beta_2 < 0$	Negative bias	Positive bias

λύση αντικατάστασης (proxy / προσεγγιστική μεταβλητή)

$$\log(wage) = \beta_0 + \beta_1 educ + \beta_2 exper + \beta_3 abil + u. \quad [9.9]$$

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3^* + u. \quad [9.10]$$

$$x_3^* = \delta_0 + \delta_3 x_3 + v_3, \quad [9.11]$$

x^* μη παρατηρούμενη / παραλειπόμενη μεταβλητή
 x παρατηρούμενη, προσεγγιστική μεταβλητή

Λύση αντικατάστασης: παλινδρομούμε y , σε $x_{1,2,3}$

Ικανοποιητική λύση όταν:

1. Εξωγένεια στην αρχική παλινδρόμηση [9.10]

2. $\text{Corr}(u, x_3) = 0$

[x_3 είναι άχρηστο αν παρατηρούσαμε το x_3^*]

3. $\text{Corr}(v_i, x_i) = 0 \quad i=1,2,3$

$\Rightarrow E(x_3^* | x_1, x_2, x_3) = E(x_3^* | x_3) = \delta_0 + \delta_3 x_3.$

[9.13]

$\pi\chi$

$E(\text{abil} | \text{educ}, \text{exper}, IQ) = E(\text{abil} | IQ) = \delta_0 + \delta_3 IQ.$

επειδή:

$$y = (\beta_0 + \beta_3 \delta_0) + \beta_1 x_1 + \beta_2 x_2 + \beta_3 \delta_3 x_3 + u + \beta_3 v_3.$$

Ή ισοδύναμα:

$$y = \alpha_0 + \beta_1 x_1 + \beta_2 x_2 + \alpha_3 x_3 + e,$$

TABLE 9.2 Dependent Variable: $\log(\text{wage})$

Independent Variables	(1)	(2)	(3)
<i>educ</i>	.065 (.006)	.054 (.007)	.018 (.041)
<i>exper</i>	.014 (.003)	.014 (.003)	.014 (.003)
<i>tenure</i>	.012 (.002)	.011 (.002)	.011 (.002)
<i>married</i>	.199 (.039)	.200 (.039)	.201 (.039)
<i>south</i>	-.091 (.026)	-.080 (.026)	-.080 (.026)
<i>urban</i>	.184 (.027)	.182 (.027)	.184 (.027)
<i>black</i>	-.188 (.038)	-.143 (.039)	-.147 (.040)
<i>IQ</i>	—	.0036 (.0010)	-.0009 (.0052)
<i>educ·IQ</i>	—	—	.00034 (.00038)
<i>intercept</i>	5.395 (.113)	5.176 (.128)	5.648 (.546)
Observations	935	935	935
<i>R</i> -squared	.253	.263	.263

TABLE 9.3 Dependent Variable: $\log(\text{crmrt}_{87})$

Independent Variables	(1)	(2)
unem_{87}	-.029 (.032)	.009 (.020)
$\log(\text{lawexp}_{87})$.203 (.173)	-.140 (.109)
$\log(\text{crmrt}_{82})$	—	1.194 (.132)
<i>intercept</i>	3.34 (1.25)	.076 (.821)
Observations	46	46
R-squared	.057	.680

λύση με μεταβλητές εργαλείο

$$y = \beta_0 + \beta_1 x + u, \quad [15.2]$$

$$\text{Cov}(x, u) \neq 0. \quad [15.3]$$

$$\log(\text{wage}) = \beta_0 + \beta_1 \text{educ} + \beta_2 \text{abil} + e,$$

$$\log(\text{wage}) = \beta_0 + \beta_1 \text{educ} + u, \quad [15.1]$$

Υποθέσεις για μεταβλητή εργαλείο z

1. Εξωγένεια εργαλείου

$$\text{Cov}(z,u) = 0;$$

[15.4]

2. Σχετικότητα εργαλείου

$$\text{Cov}(z,x) \neq 0.$$

[15.5]

πχ IQ ; εκπαίδευση μητέρας; αριθμός αδερφιών

Εκτιμητής με μεταβλητή εργαλείω

$$\hat{\beta}_1 = \frac{\sum_{i=1}^n (z_i - \bar{z})(y_i - \bar{y})}{\sum_{i=1}^n (z_i - \bar{z})(x_i - \bar{x})}.$$

[15.10]

Σύγκριση με OLS

$$\widehat{\log(wage)} = -.185 + .109 educ$$

(.185) (.014)

[15.15]

$$n = 428, R^2 = .118.$$

$$\widehat{educ} = 10.24 + .269 fatheduc$$

(.28) (.029)

[15.16]

$$n = 428, R^2 = .173.$$

$$\widehat{\log(wage)} = .441 + .059 educ$$

(.446) (.035)

[15.17]

$$n = 428, R^2 = .093.$$