2.2) Measurement Error

Vitor Kamada

August 2019



1/10

Vitor Kamada Eco 7110 Econometrics II August 2019

Reference

Wooldridge (2010). **Econometric Analysis of Cross Section and Panel Data.** Ch 4.4

https://ebookcentral.proquest.com/lib/wayne/detail.action?docID = 33391968 and the state of th

2 / 10

Vitor Kamada Eco 7110 Econometrics II August 2019

Measurement Error in the Dependent Variable

$$y^* = \beta_0 + \beta_1 x_1 + ... + \beta_k x_k + v$$
 $e_0 = y - y^*$
 $y = \beta_0 + \beta_1 x_1 + ... + \beta_k x_k + v + e_0$
 $Var(v + e_0) = \sigma_v^2 + \sigma_0^2 > \sigma_v^2$

◆ロト ◆個ト ◆ 恵ト ◆ 恵 ・ かへで

Vitor Kamada Eco 7110 Econometrics II August 2019 3 / 10

Measurement Error in Firm Scrap Rates

$$egin{aligned} &log(scrap^*) = eta_0 + eta_1 g rant + v \ &log(scrap) = log(scrap^*) + e_0 \ &log(scrap) = eta_0 + eta_1 g rant + v + e_0 \end{aligned}$$

If a firm receiving a grant is more likely to underreport its scrap rate:

$$Cov(v + e_0, grant) < 0$$

Vitor Kamada Fco 7110 Fconometrics II August 2019 4/10

Measurement Error in an Explanatory Variable

$$y = \beta_0 + \beta_1 x_1 + ... + \beta_k x_k^* + v$$
 $e_k = x_k - x_k^*$
 $y = \beta_0 + \beta_1 x_1 + ... + \beta_k x_k + (v - \beta_k e_k)$
 $Cov(x_k, e_k) = 0$
 $Var(v - \beta_k e_k) = \sigma_v^2 + \beta_k^2 \sigma_{e_k}^2$

◆□▶ ◆□▶ ◆臺▶ ◆臺▶ ■ 9000

5 / 10

Vitor Kamada Eco 7110 Econometrics II August 2019

Classical Errors-in-Variables (CEV)

$$y = \beta_0 + \beta_1 x_1 + ... + \beta_k x_k + (v - \beta_k e_k)$$

$$Cov(x_k^*, e_k) = 0$$

$$Cov(x_k, e_k)$$

$$E(x_k e_k)$$

$$E(x_k^*e_k) + E(e_k^2) = \sigma_{e_k}^2$$

Vitor Kamada Eco 7110 Econometrics II August 2019 6/10

CEV - Special Case

$$y = \beta_0 + \beta_1 x_1 + (u - \beta_1 e_1)$$

$$plim \hat{\beta}_1 = \beta_1 + \frac{Cov(x_1, u - \beta_1 e_1)}{Var(x_1)}$$

$$= \beta_1 - \frac{\beta_1 \sigma_{e_1}^2}{\sigma_{x_1}^2 + \sigma_{e_1}^2}$$

$$= \beta_1 (1 - \frac{\sigma_{e_1}^2}{\sigma_{x_1}^2 + \sigma_{e_1}^2})$$

$$= \beta_1 (\frac{\sigma_{x_1}^2}{\sigma_{x_1}^2 + \sigma_{e_1}^2})$$

Vitor Kamada Eco 7110 Econometrics II August 2019 7/10

CEV - **General Case**

$$y = \beta_0 + \beta_1 x_1 + ... + \beta_k x_k + (v - \beta_k e_k)$$
 $x_k^* = \delta_0 + \delta_1 x_1 + ... + \delta_{k-1} x_{k-1} + r_k^*$
 $Corr(x_k^*, x_j) = 0 \text{ for } j \neq k$

$$plim\hat{\beta}_{k} = \beta_{k} \left(\frac{\sigma_{r_{k}}^{2}}{\sigma_{r_{k}}^{2} + \sigma_{e_{k}}^{2}} \right)$$

4□ > 4□ > 4□ > 4□ > 4□ > 4□ >

Measurement Error in Family Income

$$CoIGPA = \ eta_0 + eta_1 faminc^* + eta_2 hsGPA + eta_3 SAT + v$$
 $faminc = faminc^* + e_1$
 $H_0: eta_1 = 0$
Type II error

4□ > 4□ > 4□ > 4□ > 4□ > 4□

Violation of CEV

$$smoked = smoked^* + e_1$$

People who do not smoke marijuana: $smoked^* = 0$ and smoked = 0

When $smoked^* > 0$, it more likely that someone miscounts

$$Corr(smoked^*, e_1) \neq 0$$



Vitor Kamada Eco 7110 Econometrics II August 2019 10 / 10