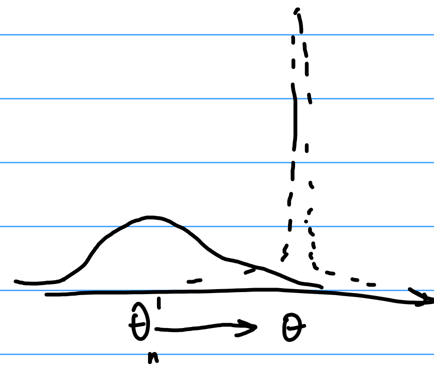
Stochastic Regressors

Endogeneous regressor $\quad \operatorname{cor}(x_i, \varepsilon_i) \neq 0$

$$\begin{tikzcd} L^s \arrow[r, "s \geq p"] & L^p \arrow[d] \\ \text{a.s.} \arrow[r] & p \arrow[r] & d \end{tikzcd}$$



in prob. : $\quad \lim\limits_{n\to\infty} \Pr(|X_n - X| > \varepsilon) = 0$

1) (asymptotic) unbiasedness

2) $\operatorname{Var} \to 0$, $n \to \infty$

in dist : $\quad \lim\limits_{n\to\infty} F_n(x) = F(x)$

LLN: $\quad y_1, \dots, y_n$ i.i.d. $\quad E(y_1) = \mu$, $\operatorname{Var}(y_1) = \sigma^2 < \infty$

$$\bar{y} \xrightarrow{P} \mu$$

CLT : $\quad y_1, \dots, y_n$ i.i.d $\quad E(y_1) = \mu$, $\operatorname{Var}(y_1) = \sigma^2 < \infty$

$$\frac{\sqrt{n}(\bar{y} - \mu)}{\sigma} \xrightarrow{d} N(0,1)$$

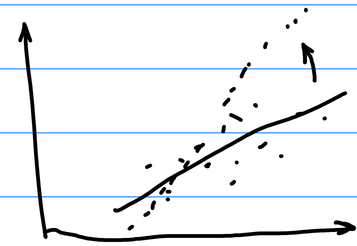
Slutsky: $\quad x_n \xrightarrow{P} a \quad \Rightarrow \quad g(x_n) \xrightarrow{P} g(a)$

$$\left.\begin{array}{l} X_n \xrightarrow{P} a \\ y_n \xrightarrow{P} b \end{array}\right\} \Rightarrow X_n + y_n \xrightarrow{P} a + b$$

TGM

1. $y_i = \beta_0 + \beta_1 x_{1i} + \dots + \beta_k x_{ki} + \varepsilon_i$
2. $\{(x_{1i}, \dots, x_{ki}, y_i),\ i = \overline{1,n}\}$ i.i.d.
3. $E(x_{ji}^4) < \infty$, $j = \overline{1,k}$; $E(y_i^4) < \infty$

   $\Leftrightarrow$ no outliers



4. $E(\varepsilon_i \mid x_{1i}, \dots, x_{ki}) = 0$

   $\hookrightarrow E(\varepsilon_i) = 0$

   $\hookrightarrow cov(\varepsilon_i, x_i) = 0$ $\leftarrow$

5. $rank(X) = k+1$ w.p. 1

1–5 $\Rightarrow$ $\hat{\beta}_{OLS}$ – consistent & asympt. normal

1. $x_i \perp \varepsilon_i$ $\Rightarrow$ unbiased, consistent
2. $cov(\varepsilon_i, x_i) = 0$ $\Rightarrow$ biased, consistent (conditionally unbiased)
3. $cov(\varepsilon_i, x_i) \neq 0$ $\Rightarrow$ inconsistent

Col. 1. : $\hat{\text{var}}(x) \xrightarrow{p} \text{Var}(x_i)$

$\hat{\text{cov}}(x,y) \xrightarrow{p} \text{Cov}(x_i, y_i)$

$$\text{cov}(x_i, y_i) = E(x_i y_i) - E(x_i)E(y_i)$$

$$\hat{\text{cov}}(x,y) = \overline{xy} - \bar{x}\cdot\bar{y}$$

$$\left.\begin{array}{l} \bar{x} \xrightarrow{p} E(x_i) \\ \bar{y} \xrightarrow{p} E(y_i) \\ \overline{xy} \xrightarrow{p} E(x_i y_i) \end{array}\right\} \text{by LLN}$$

$$\left.\begin{array}{l} \bar{x}\cdot\bar{y} \xrightarrow{p} E(x_i)E(y_i) \\ \overline{xy} - \bar{x}\cdot\bar{y} \xrightarrow{p} E(x_i y_i) - E(x_i)E(y_i) \end{array}\right\} \text{by Slutsky thm.}$$

Prob. 1. $y_i = \alpha + \beta x_i + \varepsilon_i$

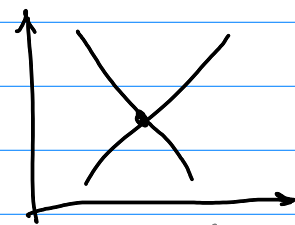
$$\text{plim}\ \hat{\beta} = \text{plim} \frac{\sum(x_i-\bar{x})(y_i-\bar{y})\,1/N}{\sum(x_i-\bar{x})^2\ \ 1/N} = \frac{\text{plim}\ \hat{\text{cov}}(x,y)}{\text{plim}\ \hat{\text{var}}(x)} =$$

$$= \frac{\text{cov}(x_i, y_i)}{\text{Var}(x_i)} = \frac{\text{cov}(x_i, \alpha+\beta x_i+\varepsilon_i)}{\text{Var}(x_i)} = \beta\frac{\cancel{\text{cov}(x_i, x_i)}}{\cancel{\text{Var}(x_i)}} + \frac{\text{cov}(x_i, \varepsilon_i)}{\text{Var}(x_i)} =$$

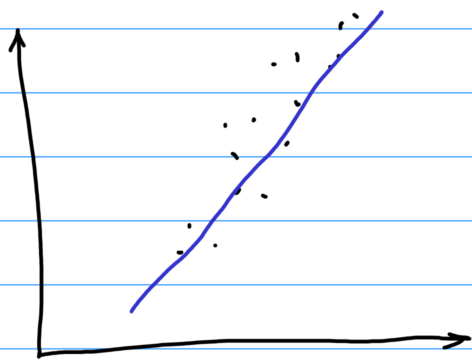
$$= \beta + \underbrace{\frac{\overset{=0}{\text{cov}(x_i, \varepsilon_i)}}{\text{Var}(x_i)}}_{\text{bias}=0} = \beta \quad \Rightarrow \quad \text{consistent}$$

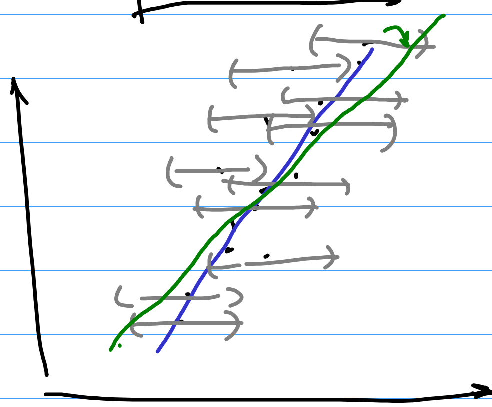
Endogeneity $\rightarrow$ omitted variable
$\rightarrow$ measurement error (for X)
$\rightarrow$ simultaneity ($Y \leftarrow X$ & $Y \rightarrow X$)

P
Q



Problem 2.





True model: $y_i = \beta_1 + \beta_2 X_i^*$ $\leftarrow$ True X

$$y_i = \beta_1 + \beta_2 (X_i - \varepsilon_i)$$
$$= \beta_1 + \beta_2 X_i - \beta_2 \varepsilon_i$$

$X_i = X_i^* + \varepsilon_i$

$\nearrow$ measured X (data)

$\nwarrow$ meas. error

$$u_i = -\beta_2 \varepsilon_i$$

Est model: $y_i = \beta_1 + \beta_2 \cdot X_i + u_i$

$$\hat{\beta}_2 \xrightarrow{p} \beta_2 + \frac{\operatorname{cov}(X_i, u_i)}{\operatorname{Var}(X_i)} = \beta_2 + \frac{\operatorname{Cov}(X_i^* + \varepsilon_i, -\beta_2 \varepsilon_i)}{\operatorname{Var}(X_i^* + \varepsilon_i)} =$$

$$= \beta_2 - \beta_2 \frac{\sigma_\varepsilon^2}{\sigma_{X^*}^2 + \sigma_\varepsilon^2} = \beta_2 \cdot \underbrace{\frac{\sigma_{X^*}^2}{\sigma_{X^*}^2 + \sigma_\varepsilon^2}}_{\text{bias}} \Rightarrow \text{inconsistent}$$

$\hat{\beta}_2$ biased towards zero $\quad |\text{bias}| < 1$