

$$y = X\beta + \varepsilon$$

i can use $\hat{\beta}_{OLS} = (X'X)^{-1}X'y$

(base one the 2 potential estimator)

or $\hat{\beta}_{TSLS} = (X'X)^{-1}X'y$

$$\text{Var } \hat{\beta}_{OLS} = \sigma_\varepsilon^2 (X'X)^{-1}$$

usually i want to compare them

$$\text{Var } \hat{\beta}_{TSLS} = \sigma_\varepsilon^2 (\tilde{X}'\tilde{X})^{-1}$$

$(\beta_{IV} - \hat{\beta}_{OLS})$ the test is based on this difference

ENDOGENEITY TEST

if the difference is = 0 then there is not endogeneity

$$H_0: \hat{\beta}_{IV} - \hat{\beta}_{OLS} = 0$$

$$H_1: \hat{\beta}_{IV} - \hat{\beta}_{OLS} \neq 0$$

HAUSMAN

$$(\hat{\beta}_{TSLS} - \hat{\beta}_{OLS})' \left[\text{Avar}(\hat{\beta}_{TSLS} - \hat{\beta}_{OLS}) \right]^{-1} (\hat{\beta}_{TSLS} - \hat{\beta}_{OLS})$$

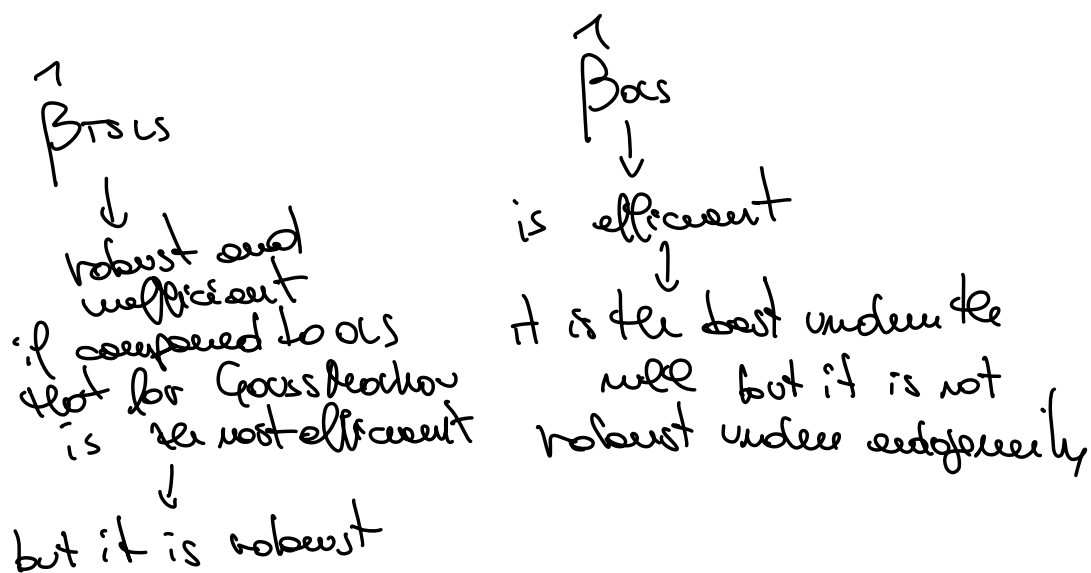
1×4 4×4 4×1
 Asymptotic

this converges in distribution to $\chi^2(k)$

only the coefficients associated with endogenous variables change

HANSEN RESULT FOR THE ASYMPTOTIC VARIANCE
OF THE DIFFERENCE

$$\text{Cov}(\hat{\beta}_{OLS}, \hat{\beta}_{IV} - \hat{\beta}_{OLS}) = 0$$



$$\text{Cov}(\hat{\beta}_{OLS}, \hat{\beta}_{IV} - \hat{\beta}_{OLS}) = 0$$

$$\text{Cov}(\hat{\beta}_{OLS}, \hat{\beta}_{IV}) - \text{Var}(\hat{\beta}_{OLS}) = 0$$

$$\text{Cov}(\hat{\beta}_{OLS}, \hat{\beta}_{IV}) = \text{Var}(\hat{\beta}_{OLS})$$

variance of the most
efficient estimator

Now $AVAR(\hat{\beta}_{TSLS} - \hat{\beta}_{OLS})$

$$= AVAR(\hat{\beta}_{IV}) + AVAR(\hat{\beta}_{OLS}) - 2Cov(\hat{\beta}_{IV}, \hat{\beta}_{OLS})$$

$$= AVAR(\hat{\beta}_{IV}) + AVAR(\hat{\beta}_{OLS}) - 2AVAR(\hat{\beta}_{OLS})$$

$$= AVAR(\hat{\beta}_{IV}) - AVAR(\hat{\beta}_{OLS})$$

$$H = (\hat{\beta}_{IV} - \hat{\beta}_{OLS})' \left[\sigma_{\varepsilon}^2 (\hat{X}'\hat{X})^{-1} - \sigma_{\varepsilon}^2 (X'X)^{-1} \right]^{-1} (\hat{\beta}_{IV} - \hat{\beta}_{OLS})$$

↓
the problem is
that a lot of

β does not
change, the

exogenous variable

being to the same
estimates in iv and ols

↓
this quantity is
positive by Gauss Markov
because $\hat{\beta}_{OLS}$ is efficient

CONTROL
REGRESSION

$$\begin{cases} y = x_1 \beta_1 + x_2 \beta_2 + \varepsilon \\ x_1 = z_1 \gamma_1 + x_2 \gamma_2 + u \\ x_2 = x_2 \end{cases}$$

we have seen strategy number 1

strategy n. 2 is estimating the following regression

$$y = x_1 \beta_1 + x_2 \beta_2 + \alpha \hat{u} + \varepsilon$$

add extra control that is
the residual from the first
equation