Specification Testing

(1) Relevance

$$\text{plim}_{n \to \infty} \frac{1}{n} W^T X = S_{W^T X}$$

Instruments must be correlated with X.

This shows up in two ways

(a) IV estimate is badly biased towards OLS when the correlation between W & X is "weak". (Greene p.350)

(b) The ~~asy~~ Asymptotics are ~~math~~ badly affected and usual approximations are misleading.

Case 1

1 endogenous variable.

In the "first stage" of 2SLS
do a joint test that
the instruments are zero.

$F > 10$ indicates they
are strong enough.

$\approx$ "10% Bias"

Example:

$$\ln(wage) = \beta_1 + \beta_2\, exp + \beta_3\, exp^2 + \beta_4\, educ + u.$$

suppose education is endogenous.
and we have instruments on
Mother's Educ. and Father's Educ.

Stage 1

$$educ = \alpha_1 + \alpha_2\, exp + \alpha_3\, exp^2 + \alpha_4\, ME$$
$$+ \alpha_5\, FE + \text{residual}.$$

H_0: $\alpha_4 = \alpha_5 = 0 \Rightarrow$ Inst weak.

H_A not H_0:

Compute F-stat and

Reject weak instruments if

$F > 10$.

Case 2

Multiple Endog. Vars.
(Greene).

let $R_k^2 = \dfrac{(X^T X)_{kk}}{(\hat{X}^T \hat{X})_{kk}}$

Shea (1997)

then

$$F = \frac{R_k^2/(\ell-1)}{(1-R_k^2)/(n-\ell)}$$

you'll have one of these for each endog. variable.

Godfrey ReSTAT 1999 for details

Exogeneity (overidentification)

$$E(W^T u) = 0$$

Technically, this is accomplished by testing the overidentifying restrictions.

$$H_0: \quad y = X\beta + u \quad u \; iid(0, \sigma^2) \quad E[W^T u] = 0$$

$$H_A: \quad y = X\beta + W^* \gamma + u \quad u \; iid(0, \sigma^2) \quad E[W^T u] = 0$$

$W^*_{n \times l-k}$ and includes surplus instruments.

Rejection can occur

(1) The model in Ho is correct, but one or more instruments are correlated with u.

(2) Model in Ho is not correctly specified. W^* omitted

(1) Estimate the Model under

H_0: are all instruments W. and get $\hat{\beta}_{IV}$ and the residuals. $\hat{u}_{IV}$

(2) ~~$Z \; Z \; X_E$~~ $: W$

$\hat{u}_{IV} = \cancel{W\hat{b}}\ W\hat{b}$ + residuals

Regress $\hat{u}_{IV}$ on all inst.

$NR^2 \sim \chi^2_{\ell-k}$ if H_0 True.

Note: W includes <u>ALL</u> exogenous variables i.e., All of the X's that are exog. + all of the instruments.

Durbin - Wu - Hausman.

H_0: $y = X\beta + u \quad u \sim (0, \sigma^2 I_n) \quad E(X^T u) = 0$

H_A: $y = X\beta + u \quad u \sim (0, \sigma^2 I_n) \quad E(W^T u) = 0$

under H_0: Both OLS & IV consistent

under H_A: OLS inconsistent, IV consistent.

under H_0: $\operatorname{plim}(\hat\beta_{IV} - \hat\beta_{OLS}) = 0$

H_A: $\neq 0$

$$\hat\beta_{IV} - \hat\beta_{OLS} = (X^T P_W X)^{-1} X^T P_W y - (X^T X)^{-1} X^T y$$

$$= (X^T P_W X)^{-1} \left[X^T P_W y - (X^T P_W X)(X^T X)^{-1} X^T y \right]$$

$$= (X^T P_W X)^{-1} X^T P_W \left[I - X (X^T X)^{-1} X^T \right] y$$

$$= (X^T P_W X)^{-1} \underbrace{X^T P_W M_X}_{} y$$

Is $\downarrow$ $= 0$?

$M_X y = \hat{u}$ least squares residuals

$P_W X = \hat{X}$

Some of the variables in W are also included in X.
These are already orthogonal to $\hat{u}$.

let $X = [\, Z \quad Y \,]$

Z's are exogenous and included in W.
Y's are potentially endogenous.

$$\underset{n \times k}{X^T} = \left[\underset{k1}{Z} \;\; \underset{k2}{Y} \right]^T$$

$$\underset{k \times n}{X^T} = \begin{pmatrix} Z^T_{k1 \times n} \\ Y^T_{k2 \times n} \end{pmatrix}$$

$$X^T P_W M_X y$$

$$\begin{pmatrix} Z^T \\ Y^T \end{pmatrix} P_W M_X y$$

$$Z^T P_W M_X y = 0_{k1 \times 1} \quad k_1 \times 1$$

$$Y^T P_W M_X y = ? \quad k_2 \times 1$$

So, $y = X\beta + P_W Y \delta + u$

$H_0: \delta = 0$

$H_A: \delta \neq 0$

$F \sim F(k2, \infty)$ if ORTHOgonal.

Rejection $\Rightarrow$ Variables in Y are not exogenous.

Why?

$$y = X\beta + P_W Y \delta + u$$

$$M_X y = M_X X\beta + M_X P_W Y \delta + M_X u.$$

$$\hat{u} = M_X P_W Y \delta + \text{res}.$$

$$\hat{\delta} = \left(Y^T P_W M_X P_W Y\right)^{-1} \left(\underbrace{Y^T P_W M_X M_X y}_{=0}\right)$$

$$\hat{\delta} = 0 \quad \Rightarrow \quad Y^T P_W M_X y = 0$$

Hausman

$$\left(\hat{\beta}_{IV} - \hat{\beta}_{LS}\right)^T \left(\mathrm{Cov}(\hat{\beta}_{IV}) - \mathrm{Cov}(\hat{\beta}_{LS})\right)^{-1} \left(\hat{\beta}_{IV} - \hat{\beta}_{LS}\right)$$

$$\rightarrow \chi^2_{k2} \text{ if } H_0 \text{ True.}$$

Slightly tricky in practice

(1) diff in est cov matrices may not be pd.

(2) Most software does not produce right df