Restriction Testing in Multiple Regression

Gujarati(2003) Chapter 7, 8, 13(pp.506-514)
Specification Error

(i) Omission of an important variable

(ii) Inclusion of an irrelevant variable

Consequences of omission of an important variable in the true model.

True model: \( Y = \beta_1 + \beta_2 X_2 + \beta_3 X_3 + u \)

\[
\begin{align*}
E(\hat{\beta}_2) &= \beta_2 \\
E(\hat{\beta}_3) &= \beta_3
\end{align*}
\]

Estimated model: \( Y = \beta^*_1 + \beta^*_2 X_2 + v \)

(Omitted an important variable)

\[
\begin{align*}
E(v) &\neq 0 \\
\text{Cov}(X_2v) &\neq 0
\end{align*}
\]

Result with biased estimator: \( E(\hat{\beta^*_2}) \neq \beta_2 \)
From the estimated model:

Since \( v = \beta_3 X_3 + u \)

and (i) \( E(v) = E(\beta_3 X_3 + u) = E(\beta_3 X_3) + E(u) \)

\( = E(\beta_3 X_3) \neq 0 \)

(ii) \( \text{Cov}(X_2, v) = \text{Cov}(X_2, \beta_3 X_3 + u) \)

\( = \beta_3 \text{Cov}(X_2 X_3) + \text{Cov}(X_2 u) \)

\( = \beta_3 \text{Cov}(X_2 X_3) \neq 0 \)

These two properties violate OLS assumptions, Therefore estimators are not BLUE.

The bias = \( \beta_3 \alpha_2 \)

Where \( \alpha_2 \) is from regression: \( X_3 = \alpha_1 + \alpha_2 X_2 + u' \)

The bias exists unless: (1) the true coefficient equals zero or

(2) the included and omitted variables are uncorrelated.
Example of Omitted Variable Bias:

True Model: \[ cons_t = \beta_1 + \beta_2 aaa_{2t} + \beta_3 dpi_{3t} + e_t \]

The Model We Estimate: \[ cons_t = \beta_1 + \beta_2 aaa_{2t} + e_t \]

Our estimated model using annual data for U.S. Economy 1959-99:

\[ cons_t = 672.14 + 192.03 aaa_t \]

\[ R^2 = 0.0742 \]

\[ \bar{R}^2 = 0.0505 \]

A corrected model

\[ cons_t = 672.14 - 17.46 aaa_t + 0.927 dpi_t \]

\[ R^2 = 0.9994 \]

\[ \bar{R}^2 = 0.9994 \]
Including a non-influential variable

True: \[ Y = \beta_1^* + \beta_2^* X + u \]

Misspecified: \[ Y = \beta_1 + \beta_2 X + \beta_3 Z + v \]

\[ \underset{E(\beta_3)}{=} 0 \]

Consequences:

(i) from OLS: \[ \hat{\beta}_2^* = \frac{\sum xy}{\sum x^2}, \quad \text{var}(\hat{\beta}_2^*) = \frac{\sigma_u^2}{\sum x^2} \]

(ii) from OLS: \[ \hat{\beta}_2 = \frac{(\sum xy)(\sum z^2) - (\sum yz)(\sum xz)}{(\sum x^2)(\sum z^2) - (\sum xz)^2} \]

\[ z = (Z - \overline{Z}) \]

\[ \text{var}(\hat{\beta}_2) = \frac{\hat{\sigma}_v^2}{\sum x^2 (1 - r_{xz}^2)} \]

This variance is not minimum
The relative efficiency:

\[
\frac{\text{Var}(\hat{\beta}_2)}{\text{Var}(\hat{\beta}_2^*)} = \frac{1}{1-r^2_{xz}}
\]

Unless X and Z are uncorrelated, \( r^2_{xz} = 0 \), otherwise the \( (\hat{\beta}_2) \) is inefficient.

It indicates \( \text{var}(\hat{\beta}_2) > \text{var}(\hat{\beta}_2^*) \)

The adjusted R\(^2\) of the misspecified model will increase.
3. Testing the addition variable in the regression model

Old model: \[ Y = \beta_1 + \beta_2 X_2 + u_1 \]

Obtain \( R^2_{\text{old}} \) or \( \text{RSS}_{\text{old}} \) and/or \( \text{ESS}_{\text{old}} \)

Now consider a new variable \( X_3 \), whether it is relevant to add or not?

New model: \[ Y = \beta_1 + \beta_2 X_2 + \beta_3 X_3 + u_2 \]

Obtain \( R^2_{\text{new}} \) or \( \text{RSS}_{\text{new}} \) and \( \text{ESS}_{\text{new}} \)

\( H_0 : \beta_3 = 0, \text{ adding } X_3 \text{ is not relevant} \)

\( H_1 : \beta_3 \neq 0, \text{ adding } X_3 \text{ is relevant} \)
Steps of testing whether $X_3$ has an incremental contribution of the explanatory power in the new model.

1. Compute the F-statistic

$$F^* = \frac{(RSS_{old} - RSS_{new}) / \# \text{ of additional variables}}{RSS_{new} / n - \# \text{ of parameters in new model}}$$

2. Compare $F^*$ and $F_{c}^{(\alpha, 1, n-3)}$

3. Decision rule: If $F^* > F_{c}$ ==> reject $H_0 : \beta_3 = 0$

that means $X_2$ is a relevant variable to be added into the model.

$F^*$ can also be calculated by

$$F = \frac{(R^2_{new} - R^2_{old}) / \text{df}}{(1 - R^2_{new}) / \text{df}}$$

# of new regressors (add or drop)

n-k (in the new model)
Add an irrelevant variable $X_5$ (R): (Studenmund, pp.166)

Old model

$$RSS_{old} = 160.5929 \quad R^2_{old} = 0.986828$$
New model

H₀: add X₅ (R) is not suitable

\[ \beta_5 = 0 \]

\[
F^* = \frac{(R^2_{\text{new}} - R^2_{\text{old}}) / df}{(1 - R^2_{\text{new}}) / df} = \frac{(0.9872 - 0.9868) / 1}{(1 - 0.9872) / 39} = 1.218
\]

\[ F^*_c 0.05, 1, 39 = 4.17 \]

Since \( F^* < F^*_c \implies \) not reject \( H_0 \).
Add an relevant variable $X_3$ (YD): (Studenmund, pp.166)

Old model

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + u$$

Next:

New model:  

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + u'$$

$H_0 :$ add $X_3$ (YD) variable is not suitable, $\beta_3 = 0$
Add a relevant variable $X_3(YD)$: (Studenmund, pp.166)

New model

\[
F^* = \frac{(R^2_{\text{new}} - R^2_{\text{old}}) / df}{(1 - R^2_{\text{new}}) / df} = \frac{0.9868 - 0.9203}{1 - 0.9868} \frac{1}{44 - 4} = \frac{0.0665 \times 40}{0.0132} = 201.5
\]

\[F^c(0.05, 1, 40) = 4.08\]

Since $F^* > F^c \implies \text{reject } H_0$. 

M. Balcilar, ECON 503, EMU
Add variables in general discussion:

old: \( Y_i = \beta_1 + \beta_2 X_2 + \beta_3 X_3 + \ldots + \beta_k X_k + u_i \)

new: \( Y_i = \beta_1 + \beta_2 X_2 + \beta_3 X_3 + \ldots + \beta_k X_k + \beta_{k+1} X_{k+1} + \beta_{k+2} X_{k+2} + u_i \)

Testing the two additional variables \( X_{k+1}, X_{k+2} \), whether they are relevant or not?

F-test:

\[ H_0 : \beta_{k+1} = 0, \beta_{k+2} = 0 \]

\[ H_1 : \beta_{k+1} \neq 0, \text{ or } \beta_{k+2} \neq 0 \]

\[ F^* = \frac{(R^2_{\text{new}} - R^2_{\text{old}})}{\text{# of added regressors}} = \frac{(1 - R^2_{\text{new}})}{n-(k+2)} \]

If \( F^* > F^c \) \( \Rightarrow \) reject \( H_0 \)
Drop variables in general discussion:

old: \( Y_i = \beta_1 + \beta_2 X_2 + \beta_3 X_3 + \ldots + \beta_k X_k + \beta_{k+1} X_{k+2} + \beta_{k+2} X_{k+2} + u_i \)

new: \( Y_i = \beta_1 + \beta_2 X_2 + \beta_3 X_3 + \ldots + \beta_k X_k + u_i \)  

Restricted \( \beta_{k+1} = \beta_{k+2} = 0 \)

Test the dropping of two variables \( X_{k+1}, X_{k+2} \) whether they are relevant or not?

F-test: \( H_0: \beta_{k+1} = 0, \beta_{k+2} = 0 \)

\( H_1: \beta_{k+1} \neq 0, \text{ or } \beta_{k+2} \neq 0 \)

\[ F^* = \frac{(R^2_{\text{old}} - R^2_{\text{new}})}{\# \text{ of dropped regressors}} \]

\[ F^* = \frac{(1 - R^2_{\text{old}})}{n-(k+2)} \]

If \( F^* > F_c \) \( \implies \) reject \( H_0 \)
Justify whether to keep (or drop) $X_{k+1}$ and $X_{k+2}$ in the model

1. **Theory**: Check the sign of $\hat{\beta}_{k+1}$ and $\hat{\beta}_{k+2}$? Whether the variables are theoretically explaining the dependent variable?

2. **Overall fitness**: Check $R^2$ increase or not? Check the $F^*$-statistics increase or not?

3. Check the **t-statistics** of $X_{k+1}$ and $X_{k+2}$? Whether $t_{k+1}^*, t_{k+2}^* > 1.96$? (5% level of significance)

4. **Bias**: Check t-statistics of other variables, $X_2, \ldots, X_k$ whether they have change significantly or not?
Use the $R^2$ instead of ESS or RSS in the F-test.

Restriction Test:

$$ F = \frac{(R^2_{UR} - R^2_R) / m}{(1 - R^2_{UR}) / (n - k)}$$

$m$: # of restricted variables

$$ = \frac{(R^2_{UR} - R^2_R) / (df_{UR} - df_R)}{(1 - R^2_{UR}) / df_{UR}}$$

$k$: # of parameters (included the intercept) in the unrestricted model
Example 8.4: The demand for Chicken (Gujarati(2003) p. 272)

Old model or restricted model
\( H_0: \) No joint effect of \( X_4 \) and \( X_5 \), i.e., \( \beta_4 = \beta_5 = 0 \)
WALD TEST

Adding variables:

1. \( H_0: \) No joint effect of \( X_4 \) and \( X_5 \), i.e., \( \beta_4 = \beta_5 = 0 \)

\[
F^* = \frac{(R^2_{\text{new}} - R^2_{\text{old}}) / \# \text{ added}}{(1-R^2_{\text{new}}) / n-k} = \frac{(0.9823 - 0.9801) / 2}{(1 - 0.9823) / (23 - 5)}
\]

\[
= \frac{0.0011}{0.000983} = 1.119
\]

Since \( F^* < F^c \implies \text{not reject } H_0 \)

\( F^c_{0.05, 2, 18} = 3.55 \)
H₀: No effect of X₅, i.e., β₅ = 0

Since t*(β₅) < tₖ \Rightarrow not reject H₀

\[ F^* = \frac{(R^2_{UR} - R^2_R)}{m} \quad \text{and} \quad 1 - R^2_{UR} / n-k \]

\[ = \frac{(0.982313 - 0.981509)}{1} \quad \text{and} \quad \frac{(1-0.982313)}{(23 - 4)} \]

\[ = 0.864 \quad < F_{0.05, 1, 19} = 4.38 \quad \Rightarrow \text{not reject } H₀ \]
4. Testing partial coefficient under some restrictions:

Cobb-Douglas production

Restricted least squares:
Constant returns to scales

\[ Y = \beta_1 X_2^{\beta_2} X_3^{\beta_3} e^u \]

\[ \beta_2 + \beta_3 = 1 \]

\[ \beta_2 = 1 - \beta_3 \]

\[ \beta_3 = 1 - \beta_2 \]

Unrestricted model

Restricted model

\[ \ln Y = \beta_1 + \beta_2 \ln X_2 + \beta_3 \ln X_3 + u \]

\[ \Rightarrow \ln Y = \beta_1 + (1 - \beta_3) \ln X_2 + \beta_3 \ln X_3 + u \]

\[ \Rightarrow \ln Y = \beta_1 + \ln X_2 + \beta_3 (\ln X_3 - \ln X_2) + u \]

\[ \Rightarrow (\ln Y - \ln X_2) = \beta_1 + \beta_3 (\ln X_3 - \ln X_2) + u \]

\[ \Rightarrow \ln \left( \frac{Y}{X_2} \right) = \beta'_1 + \beta'_3 \ln \left( \frac{X_3}{X_2} \right) + u' \]

\[ Y^* = \beta'_1 + \beta'_3 X^* + u' \]
OR

=> \( \ln Y = \beta_1 + \beta_2 \ln X_2 + (1 - \beta_2) \ln X_3 + u \)

=> \( \ln Y = \beta_1 + \beta_2 \ln X_2 + \ln X_3 - \beta_2 \ln X_3 + u \)

=> \( (\ln Y - \ln X_3) = \beta_1 + \beta_2 (\ln X_2 - \ln X_3) + u \)

=> \( \ln \left( \frac{Y}{X_3} \right) = \beta''_1 + \beta''_2 \ln \left( \frac{X_2}{X_3} \right) + u'' \)

\( Y^{**} = \beta''_1 + \beta''_3 X^{**} + u'' \)
Unrestricted equation: 
\[ \ln(Y) = \beta_1 + \beta_2 \ln(X_2) + \beta_3 \ln(X_3) + u \]

Restricted equation: 
\[ \ln(Y/X_2) = \beta'_1 + \beta'_3 \ln(X_3/X_2) \]

H₀: \( \beta_2 + \beta_3 = 1 \)

\[ F = \frac{(RSS_R - RSS_{UR})/m}{(RSS_{UR})/(n - k)} \]

\( F^* = 3.78 \)

\( F^c_{(0.05,1,17)} = 4.45 \)

\[ RSS_{UR} = 0.013604 \]
\[ RSS_R = 0.016629 \]

F cannot be calculated from \( R^2 \)'s since dependent variables are different.
Unrestricted equation: 
\[ \ln Y = \beta_1 + \beta_2 \ln X_2 + \beta_3 \ln X_3 + u \]

Restricted equation: 
\[ \ln \left( \frac{Y}{X_3} \right) = \beta'_1 + \beta'_2 \ln \left( \frac{X_2}{X_3} \right) \]

\[ H_0 : \beta_2 + \beta_3 = 1 \]

\[ F = \frac{(RSS_R - RSS_{UR})/m}{RSS_{UR}/(n - k)} \]

\[ F* = 3.78 \]

Not Reject \( H_0 \)

\[ RSS_{UR} = 0.013604 \]
\[ RSS_R = 0.016629 \]

\[ F* = 3.78 \]

M. Balciar, ECON 503, EMU
Test for Restriction on parameters:

Unrestricted model: \( Y = \beta_1 + \beta_2 X_2 + \beta_3 X_3 + u \)

Restriction: \( \beta_2 = \beta_3 \) (or \( \beta_2 - \beta_3 = \theta = 0 \) or as \( \beta_2 = \theta + \beta_3 = 0 \) )

\[
\text{Compute } t = \frac{\hat{\beta}_2 - \hat{\beta}_3}{\text{se}(\hat{\beta}_2 - \hat{\beta}_3)} \quad \text{and compare to } t^c \\
\text{and follow } t\text{-test decision rule}
\]

\[
\text{se}(\hat{\beta}_2 - \hat{\beta}_3) = \sqrt{\text{var}(\hat{\beta}_2) + \text{var}(\hat{\beta}_3) - 2 \text{cov}(\hat{\beta}_2, \hat{\beta}_3)}
\]

Next rewrite the equation as:

\[
\Rightarrow Y = \beta_1 + (\theta + \beta_3)X_2 + \beta_3 X_3 + u'
\]
\[
\Rightarrow Y = \beta_1 + \theta X_2 + \beta_3 X_2 + \beta_3 X_3 + u'
\]
\[
\Rightarrow Y = \beta_1 + \theta X_2 + \beta_3 (X_2 + X_3) + u'
\]

Restricted model: \( Y = \beta_1 + \theta X_2 + \beta_3 X^*_3 + u' \)

Simply use the t-value to test whether \( \theta \) is zero or not
$H_0 : \beta_3 - \beta_4 = 0$

$$F^* = \frac{(R^2_{UR} - R^2_R) / m}{(1 - R^2_{UR}) / n - k} = 0$$

$$F^c_{(0.05, 4, 931)} = 2.37$$

$H_0 : \theta = 0$
$H_1 : \theta \neq 0$

$$t^c_{(0.05, 931)} = 1.96$$

$\Rightarrow$ not reject $H_0$
Wald, Likelihood Ratio (LR), and Lagrange Multiplier (LM) Tests

• The F restriction test we examined above is the F version of the Wald test.
• There is also a $\chi^2$ variant of the Wald test.
• Additionally, two commonly used tests exists, the LR and LM tests.
Consider J linear restrictions: \( H_0: \beta_1 = r_1, \ldots, \beta_m = r_m \)
\( H_1: \beta_1 \neq r_1, \ldots, \beta_m = r_m \)

The unrestricted model under \( H_1 \) is
\[
Y_i = \beta_1 + \beta_2 X_{2i} + \cdots + \beta_{l-1} X_{l-1,i} + \beta_l X_{li} + \cdots + \beta_m X_{mi} + \cdots + \beta_k X_{ki} + u_i
\]

Let the residual sum of squares from estimating this unrestricted model be
\[
RSS_{UR} = \sum \hat{u}_i^2
\]

In order to test these J restrictions, we estimate the same model by imposing the restriction:
\[
Y_i = \beta_1 + \beta_2 X_{2i} + \cdots + \beta_{l-1} X_{l-1,i} + r_l X_{li} + \cdots + r_m X_{mi} + \cdots + \beta_k X_{ki} + u_i^*
\]

Let the residual sum of squares from estimating this unrestricted model be
\[
RSS_R = \sum \hat{u}_{i}^{*2}
\]
Likelihood Ratio Test

Log-likelihood of the unrestricted model:

\[ L_{UR} = -\frac{n}{2} \ln(2\pi) - \ln(\sigma^2) - \frac{n}{2\sigma^2} \sum (Y_i - \beta_1 - \beta_2 X_{2i} - \cdots - \beta_{l-1} X_{l-1,i} - \beta_l X_{li} - \cdots - \beta_m X_{mi} - \cdots - \beta_k X_{ki})^2 \]

\[ = -\frac{n}{2} \ln(2\pi) - \ln(\sigma^2) - \frac{n}{2\sigma^2} \sum u_t^2 \]

Log-likelihood of the restricted model:

\[ L_R = -\frac{n}{2} \ln(2\pi) - \ln(\sigma^2) - \frac{n}{2\sigma^2} \sum (Y_i - \beta_1 - \beta_2 X_{2i} - \cdots - \beta_{l-1} X_{l-1,i} - \beta_l X_{li} - \cdots - \beta_m X_{mi} - \cdots - \beta_k X_{ki})^2 \]

\[ = -\frac{n}{2} \ln(2\pi) - \ln(\sigma^2) - \frac{n}{2\sigma^2} \sum u_t^{*2} \]
Likelihood Ratio Test

max of the log-likelihood of the unrestricted model: \( \hat{L}_{UR} \)

max of the Log-likelihood of the restricted model: \( \hat{L}_R \)

TEST STATISTICS:

\[ LR = 2(\hat{L}_{UR} - \hat{L}_R) \text{ or } LR = n \ln \left( \frac{RSS_R}{RSS_{UR}} \right) = n(\ln RSS_R - \ln RSS_{UR}) \]

DISTRIBUTION:

\[ LR \sim \chi^2(m) \]

where \( m = \) degrees of freedom = \# of restrictions

DECISION:

If \( LR > \chi^2(m) \) \rightarrow Reject H_0

If \( LR \leq \chi^2(m) \) \rightarrow Do not Reject H_0
Wald Test

RSS of the unrestricted model: $\text{RSS}_{UR}$

RSS of the restricted model: $\text{RSS}_R$

**TEST STATISTICS:**

F-variant: $F = \frac{(\text{RSS}_R - \text{RSS}_{UR})/m}{\text{RSS}_{UR}/(n-k)}$

Chi-Square-variant: $W = \frac{n(\text{RSS}_R - \text{RSS}_{UR})}{\text{RSS}_{UR}} = \frac{nm}{n-k} F$

**DISTRIBUTION:**

$F \sim F(m, n-k)$

$W \sim \chi^2(m)$

**DECISION:**

If $F > F_c$ → Reject $H_0$

If $F \leq F_c$ → Do not Reject $H_0$

If $W > \chi^2,c$ → Reject $H_0$

If $W \leq \chi^2,c$ → Do not Reject $H_0$
**LM Test**

RSS of the **unrestricted** model: $RSS_{UR}$

RSS of the **restricted** model: $RSS_R$

**TEST STATISTICS:**

$$LM = \frac{n(RSS_R - RSS_{UR})}{RSS_R}$$

**DISTRIBUTION:**

$LM \sim \chi^2(m)$

where $m$ = degrees of freedom = # of restrictions

**DECISION:**

If $LM > \chi^2_c$ \quad Reject $H_0$

If $LM \leq \chi^2_c$ \quad Do not Reject $H_0$
LM Test

Alternative method to calculate the LM test:

Step 1: Estimate the restricted model and calculate $\hat{u}_t^*$:

$$Y_i = \beta_1 + \beta_2 X_{2i} + \cdots + \beta_{l-1} X_{l-1,i} + r_1 X_{li} + \cdots + r_m X_{mi} + \cdots + \beta_k X_{ki} + u_i^*$$

Step 2: Estimate the following auxiliary regression and obtain its $R^2$:

$$\hat{u}_i^* = \alpha_1 + \alpha_2 X_{2i} + \cdots + \alpha_{l-1} X_{l-1,i} + \alpha_l X_{li} + \cdots + \alpha_m X_{mi} + \cdots + \alpha_k X_{ki} + v_i$$

Step 3: Calculate the LM statistics using the $R^2$ from Step 2:

$$LM = nR^2$$

**DISTRIBUTION:**

$$LM \sim \chi^2 (m)$$

**DECISION:**

If $LM > \chi^{2,c}$, Reject $H_0$

If $LM \leq \chi^{2,c}$, Do not Reject $H_0$
Example: LR, LM, Wald Tests

Import demand function for Turkey: 39 obs (87:Q1-95:Q4)

- PM = import price index (foreign prices) (1987=100)
- PD = wholesale price index (domestic prices) (1987=100)
- Y = real GDP at 1987 producer prices
- M = real imports (nominal)

**Unrestricted Model:**

\[
\ln M_t = \beta_1 + \beta_2 \ln Y_t + \beta_3 \ln PD_t + \beta_4 \ln PM_t + u_t
\]

**Restriction:**

\( H_0: \beta_4 = -\beta_3 \) (price homogeneity, consumers respond to domestic \((\beta_3 + \beta_4 = 0)\) and foreign price changes in the same manner)

**Restricted Model:**

\[
\ln M_t = \beta_1 + \beta_2 \ln Y_t + \beta_3 (\ln PM_t - \ln PD_t) + u_t^*
\]
LR Test

\[ LR = 2(\hat{L}_{UR} - \hat{L}_R) \] or \[ LR = n(\ln RSS_R - \ln RSS_{UR}) \]

\[ LR = 2*(21.16697-21.08045)=0.17304 \] or \[ LR = 36*(\ln(0.653447)-\ln(0.650313))=0.1730751 \]

\[ LR = 0.17304 < \chi^2_{0.05}(1) = 3.84 \]

\[ LR = 0.1730751 < \chi^2_{0.05}(1) = 3.84 \]

\[ \Rightarrow \text{Do not reject the null} \Rightarrow \text{important demand function has price homogeneity} \]

M. Balcilar, ECON 503, EMU
Wald Test

\[ F = \frac{(RSS_R - RSS_{UR})/m}{RSS_{UR}/(n-k)} = \frac{(0.653447 - 0.650313)/1}{0.650313/(36-4)} = 0.1542150 \]

\[ F = 0.1542150 < F_{0.05}(1,32) = 4.149097 \]

\[ W = \frac{n(RSS_R - RSS_{UR})}{RSS_{UR}} = \frac{36(0.653447 - 0.650313)}{0.650313} = 0.1734918 \]

\[ W = 0.1734918 < \chi^2_{0.05}(1) = 3.84 \]

\[ \rightarrow \text{ Do not reject the null } \rightarrow \text{ important demand function has price homogeneity} \]
EViews assumes \( W = \frac{(n-k)(RSS_R - RSS_{UR})}{RSS_{UR}} \) so, \( W = mF \)

when \( m = 1 \) we get \( W = F \)

At \( \alpha = 5\% \), p-value > \( \alpha \) \( \Rightarrow \) Do not reject the null
\( \Rightarrow \) important demand function has price homogeneity
LM Test

\[ LM = \frac{n(RSS_R - RSS_{UR})}{RSS_R} = \frac{36 \times (0.653447 - 0.650313)}{0.653447} \]
\[ = 0.1726598 \]

\[ LM = 0.1726598 < \chi^2_{0.05}(1) = 3.84 \]

Do not reject the null \( \Rightarrow \) important demand function has price homogeneity
LM Test

\[ LM = nR^2 \]
\[ = 36 \times 0.004795 \]
\[ = 0.17262 \]
\[ LM = 0.17262 < \chi^2_{0.05} (1) = 3.84 \]

\( \Rightarrow \) Do not reject the null  \( \Rightarrow \) important demand function has price homogeneity

A well known inequality:

\[ W \geq LR \geq LM \]

\[ 0.17350 > 0.17304 > 0.17262 \]
Exact and Large Sample Tests

- In some cases we were able to obtain a test statistic $z$ that was distributed as $N(0, 1)$. Tests based on this statistic are **exact**.
- Unfortunately, it is possible to perform exact tests only in certain special cases. One very important special case of this type arises when we test linear restrictions on the parameters of the classical normal linear model.
- The $t$ and $F$ tests we examined previously are exact only under the strong assumptions of the classical **normal** linear model.
- If the error vector were not normally distributed or not independent of the regressors, we could still compute $t$ and $F$ statistics, but they would not actually follow their namesake distributions in finite samples.
- However, like a great many test statistics in econometrics which do not follow any known distribution exactly, they would in many cases approximately follow known distributions in large samples. In such cases, we can perform what are called large-sample tests or asymptotic tests, using the approximate distributions to compute $P$ values or critical values.
- Wald, LR, and LM tests are large sample (asymptotic) test.
Summary: Wald, LR, LM Tests

Wald Test:
F-variant: 
\[ F = \frac{(RSS_R - RSS_{UR}) / m}{RSS_{UR} / (n - k)} \]
Chi-Square-variant: 
\[ W = \frac{n(RSS_R - RSS_{UR})}{RSS_{UR}} \]

LR Test: 
\[ LR = 2(\hat{L}_{UR} - \hat{L}_R) \text{ or } LR = n \ln \left( \frac{RSS_R}{RSS_{UR}} \right) = n(\ln RSS_R - \ln RSS_{UR}) \]

LM Test: 
\[ LM = \frac{n(RSS_R - RSS_{UR})}{RSS_R} \]

Relations among F, W, LR, LM:
\[ W = \frac{nm}{n - k} F \]
\[ LR = n \ln \left( 1 + \frac{m}{n - k} F \right) \]
\[ LM = \frac{nm}{(n - k)F^{-1} + m} \]

A well known inequality:
\[ W \geq LR \geq LM \]
"Chow Test" - structural stability Test

H₀ : no structural change (parameters are same for both subsamples)
H₁ : yes

Procedures:
1. Divide the sample of n observations into two groups.
   - group1 consisting of first n₁ obs.
   - group2 consisting of the remaining n₂ = n - n₁ obs.
2. Run OLS on two sub-sample groups separately and obtain the $RSS_1$, and $RSS_2$

3. Run OLS on the whole sample (N) and obtain the restricted $RSS_R$

4. Compute
   \[
   F^* = \frac{(RSS_R - RSS_1 - RSS_2) / k}{(RSS_1 + RSS_2) / (n-2k)}
   \]

5. Compute $F^*$ and $F^c_{\alpha, k, N-2k}$
   
   If $F^* > F^c$ \(\Rightarrow\) reject $H_0$

   It means that there is a structural change in the sample.
Structural stability: CHOW TEST

1970 - 1995

SAVINGS

INCOME
Scatter plot of Income and Savings

1970 - 1981

1982 - 1995
Structural stability: $H_0: \text{Var}(u_1) = \text{Var}(u_2) = \sigma^2$

![Image of statistical software output]

**Whole sample**

- **Dependent Variable:** SAVINGS
- **Method:** Least Squares
- **Date:** 01/28/03, **Time:** 18:28
- **Sample:** 1970 1995
- **Included observations:** 26

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Std. Error</th>
<th>t-Statistic</th>
<th>Prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>62.42267</td>
<td>12.76075</td>
<td>4.891772</td>
<td>0.0001</td>
</tr>
<tr>
<td>INCOME</td>
<td>0.037679</td>
<td>0.004237</td>
<td>8.893776</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

- **R-squared:** 0.767215 (Mean dependent var: 162.0885)
- **Adjusted R-squared:** 0.757515 (S.D. dependent var: 63.20446)
- **S.E. of regression:** 31.12361 (Akaike info criterion: 9.787614)
- **Sum squared resid:** 23248.30 (Schwarz criterion: 9.884390)
- **Log likelihood:** -125.2390 (F-statistic: 79.09925)
- **Durbin-Watson stat:** 0.859717 (Prob(F-statistic): 0.000000)
The equation for the model is:

\[ Y = \alpha_1 + \alpha_2 X + u_1 \]

The sub-sample is from 1970 to 1981, and the included observations are 12.

### Regression Output

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Std. Error</th>
<th>t-Statistic</th>
<th>Prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>1.016117</td>
<td>11.63771</td>
<td>0.087313</td>
<td>0.9321</td>
</tr>
<tr>
<td>INCOME</td>
<td>0.080332</td>
<td>0.008367</td>
<td>9.601576</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

### Summary Statistics

- **R-squared**: 0.902143
- **Adjusted R-squared**: 0.892358
- **S.E. of regression**: 13.36051
- **Sum squared resid**: 1785.032
- **Log likelihood**: -47.04097
- **Durbin-Watson stat**: 0.864230
- **Mean dependent var**: 106.4417
- **S.D. dependent var**: 40.72222
- **Akaike info criterion**: 8.173495
- **Schwarz criterion**: 8.254313
- **F-statistic**: 92.19026
- **Prob(F-statistic)**: 0.000002
\[ Y = \beta_1 + \beta_2 X + u_2 \]

Sub-sample \( n_2 \)

- RSS\(_2\)
- Included observations: 14
- Dependent Variable: SAVINGS
- Method: Least Squares
- Date: 01/28/03  Time: 18:31
- Sample: 1982 1995
- Sum squared resid: 10005.22
- Durbin-Watson stat: 1.786588
Empirical Results:

<table>
<thead>
<tr>
<th>Dep. variable</th>
<th>Constant</th>
<th>Indep. V</th>
<th>$R^2$</th>
<th>SEE</th>
<th>RSS</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>Y (70-95)</td>
<td>624226</td>
<td>0.0376</td>
<td>0.7672</td>
<td>31.12</td>
<td>23248.3</td>
<td>26</td>
</tr>
<tr>
<td></td>
<td>(4.89)</td>
<td>(8.89)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Y (70-81)</td>
<td>1.0161</td>
<td>0.0803</td>
<td>0.9021</td>
<td>13.36</td>
<td>1785.03</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>(0.08)</td>
<td>(9.60)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Y (82-95)</td>
<td>153.494</td>
<td>0.0148</td>
<td>0.2071</td>
<td>28.87</td>
<td>10005.2</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>(4.69)</td>
<td>(1.77)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\[
F^* = \frac{(\text{RSS}_R - \text{RSS}_1 - \text{RSS}_2) / k}{(\text{RSS}_1 + \text{RSS}_2) / (n-2k)}
\]

= \frac{(23248.3 - 1785.03 -10005.2)/ 2}{(1785.03 +10005.2)/(22)}

\[
F^* = 10.69
\]

Conclusion: $F^* > F_c \Rightarrow$ reject $H_0$

$F_{0.01}^c = 5.72; \; F_{0.10}^c = 2.56; \; F_{0.05}^c = 3.44$
GENR
dummy = 0  for 1970-1981
dummy = 1  for 1982-1995
<table>
<thead>
<tr>
<th>year</th>
<th>Dummy</th>
</tr>
</thead>
<tbody>
<tr>
<td>1970</td>
<td>0</td>
</tr>
<tr>
<td>1971</td>
<td>0</td>
</tr>
<tr>
<td>1972</td>
<td>0</td>
</tr>
<tr>
<td>1973</td>
<td>0</td>
</tr>
<tr>
<td>1974</td>
<td>0</td>
</tr>
<tr>
<td>1975</td>
<td>0</td>
</tr>
<tr>
<td>1976</td>
<td>0</td>
</tr>
<tr>
<td>1977</td>
<td>0</td>
</tr>
<tr>
<td>1978</td>
<td>0</td>
</tr>
<tr>
<td>1979</td>
<td>0</td>
</tr>
<tr>
<td>1980</td>
<td>0</td>
</tr>
<tr>
<td>1981</td>
<td>0</td>
</tr>
<tr>
<td>1982</td>
<td>1</td>
</tr>
<tr>
<td>1983</td>
<td>1</td>
</tr>
<tr>
<td>1984</td>
<td>1</td>
</tr>
<tr>
<td>1985</td>
<td>1</td>
</tr>
<tr>
<td>1986</td>
<td>1</td>
</tr>
<tr>
<td>1987</td>
<td>1</td>
</tr>
</tbody>
</table>
Using the dummy variable to identify the structural instability

Check the t-statistics

Generate a dummy series “DUMMY”
Equation Specification:

Dependent variable followed by list of regressors including ARMA and PDL terms, OR an explicit equation like \( Y = c(1) + c(2)X \).

\[
\text{savings} \quad c \quad \text{dummy} \quad \text{income} \quad \text{dummy}^{\times\text{income}}
\]

Estimation Settings:

Method: LS - Least Squares (NLS and ARMA)

Sample: 1946 1963
Read the Results from the dummy regression:

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Std. Error</th>
<th>t-Statistic</th>
<th>Prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>1.016117</td>
<td>20.16483</td>
<td>0.050391</td>
<td>0.9603</td>
</tr>
<tr>
<td>DUMMY</td>
<td>152.4786</td>
<td>33.08237</td>
<td>4.609058</td>
<td>0.0001</td>
</tr>
<tr>
<td>INCOME</td>
<td>0.080332</td>
<td>0.014497</td>
<td>5.541347</td>
<td>0.0000</td>
</tr>
<tr>
<td>DUMMY*INCOME</td>
<td>-0.065469</td>
<td>0.015982</td>
<td>4.096340</td>
<td>0.0005</td>
</tr>
</tbody>
</table>

Slope change?

Intercept change?