Lecture 5
Testing in the CLM

Brooks (4^{th} edition): Chapters 3 & 4

 1

Review: Bootstrapping in the CLM Var[b]

• We use a bootstrap to estimate **b**, Var[**b**], and t-stats. We can also compute C.I. for **b**.

• Steps to bootstrap **b** in the CLM:

1. Estimate CLM using full sample (of size *T*) ⇒ get **b**
2. Repeat *B* times:
   - Draw *T* observations from the sample, *with replacement*
   - Estimate **β** with mean of **b**(r).
3. Estimate variance with
   $\mathbf{V}_{\mathbf{boot}} = (1/B)\ [\mathbf{b}(r) - \mathbf{b}][\mathbf{b}(r) - \mathbf{b}]'$
   (Square root along the diagonal of $\mathbf{V}_{\mathbf{boot}}$ gives SE[**b**(r)]).
4. Estimate t-stats with
   t = meam(**b**(r)/SE[**b**(r)]

Review: Bootstrapping in the CLM

• Comparing OLS and Bootstrap Estimation for the FF 3-factor model for IBM returns:

	OLS		Bootstrap		Bias
	Coeff. (1)	S.E.	Coeff. (2)	S.E.	(2)-(1)
x	-0.00509	0.00249	-0.00501	0.00249	8.0765e-05
xMkt_RF	**0.90829**	**0.05672**	**0.90684**	**0.06132**	-0.0014571
xSMB	**-0.21246**	**0.08411**	**-0.21245**	**0.11080**	1.9914e-06
xHML	**-0.17150**	**0.08468**	**-0.17099**	**0.09730**	0.0005133

• Higher SE for the bootstrap: More conservative tests (less rejections of H_0).When in doubt, always use more conservative tests.

Review – OLS Assumptions

• CLM Assumptions

(**A1**) DGP: $\mathbf{y} = \mathbf{X}\,\boldsymbol{\beta} + \boldsymbol{\varepsilon}$ is correctly specified.

(**A2**) $E[\boldsymbol{\varepsilon}|\mathbf{X}] = 0$

(**A3**) $Var[\boldsymbol{\varepsilon}|\mathbf{X}] = \sigma^2\,\mathbf{I}_T$

(**A4**) $\mathbf{X}$ has full column rank –rank($\mathbf{X}$)=k-, where $T \geq k$.

Q: What happens when we impose to the DGP (**A1**) a linear restrictions, $\mathbf{R}\boldsymbol{\beta} = \mathbf{q}$?

A: We get a restricted estimator $\Rightarrow \mathbf{b}^*$

Q: How do we test joint restrictions in the context of OLS?

A: We use Wald tests & F-tests.

Review: OLS Subject to Linear Restrictions

• Restrictions: Theory imposes certain restrictions on parameters and provide the foundation of several tests. In this Lecture, we only consider linear restrictions, written as $\mathbf{R}\beta = \mathbf{q}$.

Dimensions:

$\mathbf{R}$: Jxk - J = # of restrictions & k = # of pars.

β: kx1

$\mathbf{q}$: kx1

• We consider the following restrictions:

(1) Dropping variables from model ($\beta_{SMB} = 0$).

(2) Adding up conditions ($\beta_{SMB} + \beta_{HML} = 1$).

(3) Equality restrictions ($\beta_{SMB} = \beta_{HML} = 0$).

Review: OLS Subject to Linear Restrictions

• We have a programming problem:

Minimize wrt β $L^* = (\boldsymbol{y} - \mathbf{X}\beta)'(\boldsymbol{y} - \mathbf{X}\beta)$ s.t. $\mathbf{R}\beta = \mathbf{q}$

• The Lagrangean approach:

Min $_{\mathbf{b},\lambda}$ $\{ L^* = (\boldsymbol{y} - \mathbf{X}\beta)'(\boldsymbol{y} - \mathbf{X}\beta) + 2\,\lambda\,(\mathbf{R}\beta - \mathbf{q}) \}$

• After (a lot of algebra) we get:

Restricted LS estimator: $\mathbf{b}^* = \mathbf{b} - (\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'[\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}']^{-1}(\mathbf{R}\mathbf{b} - \mathbf{q})$

$= \mathbf{b}$ + correction

• Properties:

1. Unbiased?

- Yes, if Theory is correct: $E[\mathbf{b}^* | \mathbf{X}] = E[\mathbf{b} | \mathbf{X}] = \beta$
- No, if Theory is incorrect: $E[\mathbf{b}^* | \mathbf{X}] \neq \beta$

Restricted LS – R^{*2}

- Properties:

1. Unbiased? Yes, if Theory is correct! $E[\mathbf{b}^* | \mathbf{X}] = \boldsymbol{\beta}$

2. Efficiency? Yes. $\text{Var}[\mathbf{b}^* | \mathbf{X}] < \text{Var}[\mathbf{b} | \mathbf{X}]$

3. A biased $\mathbf{b}^*$ may be more "precise," using metric MSE (=RSS/T)
   MSE =RSS/T = Variance + Squared Bias

4. We can show that RSS never decreases with restrictions:
   $\mathbf{e}'\mathbf{e} = (\boldsymbol{y} - \mathbf{Xb})'(\boldsymbol{y} - \mathbf{Xb}) \leq \mathbf{e}^{*\prime}\mathbf{e}^* = (\boldsymbol{y} - \mathbf{Xb}^*)'(\boldsymbol{y} - \mathbf{Xb}^*)$

   $\Rightarrow$ Restrictions cannot increase R^2 $\Rightarrow R^2 \geq R^{2*}$

Wald Statistic

- Most of our test statistics, including joint tests, are Wald statistics.
  Wald = normalized distance measure.

One parameter: $t_k = \dfrac{b_k - \beta_k^0}{\text{SE}[b_k]}$ = distance/unit

More than one parameter.
Let $\mathbf{z}$ = (random vector – hypothesized value) be the distance
$W = \mathbf{z}' [\text{Var}(\mathbf{z})]^{-1} \mathbf{z}$ -a quadratic form, produces a number

Example: Let $\mathbf{z} = \mathbf{Rb} - \mathbf{q}$, which under (**A5**) & H_0: $\mathbf{R}\boldsymbol{\beta} = \mathbf{q}$
$\mathbf{z} \sim N(\mathbf{0}, \text{Var}[\mathbf{z}])$ $\text{Var}[\mathbf{z}] = \mathbf{R}\, [\text{Var}[\mathbf{b} | \mathbf{X}]]^{-1}\, \mathbf{R}'$

Then, if H_0 is correct, W should be a small number, ideally close to zero. A large value would be evidence against H_0.

We need the distribution of W to determine how "far" is from zero.

Wald Statistic

- Distribution of W? We have a quadratic form.
  - If **z** is normal and σ^2 known, $W \sim \chi^2_{rank[Var(z)]}$
  - If **z** is normal and σ^2 unknown, which we estimate with $s^2 = \mathbf{e'e}/(T - k)$, then $W \sim F$
  - If **z** is not normal and we use s^2 to estimate the unknown σ^2, we rely on asymptotic theory, then $W \xrightarrow{d} \chi^2_{rank[Var(z)]}$


Abraham Wald (1902–1950, Hungary)

Testing H_0: $\mathbf{R}\boldsymbol{\beta} - \mathbf{q} = \mathbf{0}$

- Q: Is $\mathbf{Rb} - \mathbf{q}$ close to $\mathbf{0}$? Two different approaches to this question.

Approach (1): We base the answer on the discrepancy vector:

$$\mathbf{m} = \mathbf{Rb} - \mathbf{q}.$$

Then, we construct a Wald statistic:

$$W = \mathbf{m'}\,(\text{Var}[\mathbf{m}\,|\,\mathbf{X}])^{-1}\,\mathbf{m}$$

to test if **m** is different from 0.

Approach (2): We base the answer on a model loss of fit when restrictions are imposed: RSS must increase and R^2 must go down. Then, we construct an F test to check if the unrestricted RSS (RSS_U) is different from the restricted RSS (RSS_R).

Review: Testing H_0: $R\beta - q = 0$ with a Wald Test

Approach (1): Test H_0 with $W = \mathbf{m}' \, (Var[\mathbf{m}|\mathbf{X}])^{-1} \, \mathbf{m}$

Based on unrestricted OLS estimation we compute:

$$\mathbf{m} = \mathbf{Rb} - \mathbf{q} \qquad \text{(under (A5) \& } H_0\text{: } \mathbf{m} \sim N(\mathbf{0}, Var[\mathbf{m}]))$$
$$Var[\mathbf{m}|\mathbf{X}] = \mathbf{R}\,[\sigma^2(\mathbf{X}'\mathbf{X})^{-1}]\,\mathbf{R}'$$

Then, we compute the Wald statistic:

$$W = (\mathbf{Rb} - \mathbf{q})' \, \{\mathbf{R}[\sigma^2(\mathbf{X}'\mathbf{X})^{-1}]\mathbf{R}'\}^{-1} \, (\mathbf{Rb} - \mathbf{q})$$

Under H_0 and assuming (**A5**) & estimating σ^2 with $s^2 = \mathbf{e}'\mathbf{e}/(T - k)$:

$$W^* = (\mathbf{Rb} - \mathbf{q})' \, \{\mathbf{R}[s^2(\mathbf{X}'\mathbf{X})^{-1}]\mathbf{R}\}^{-1} \, (\mathbf{Rb} - \mathbf{q})$$
$$F = W^*/J \sim F_{J,T-k}.$$

If (**A5**) is not assumed, the results are only asymptotic: $J\,F \xrightarrow{d} \chi^2_J$

Review: Testing H_0: $R\beta - q = 0$ with a Wald Test

- Under H_0 and assuming (**A5**) & estimating σ^2 with $s^2 = \mathbf{e}'\mathbf{e}/(T - k)$:

$$W^* = (\mathbf{Rb} - \mathbf{q})' \, \{\mathbf{R}[s^2(\mathbf{X}'\mathbf{X})^{-1}]\mathbf{R}\}^{-1} \, (\mathbf{Rb} - \mathbf{q})$$
$$F = W^*/J \sim F_{J,T-k}.$$

<u>Technical note</u>: Why the F distribution?

The F-distribution is a ratio of two independent χ^2_J and χ^2_T RV divided by their degrees of freedom: $F = \frac{\chi^2_J / J}{\chi^2_T / T} \sim F_{J,T}$

(1) Numerator: $W = (\mathbf{Rb} - \mathbf{q})' \, \{\mathbf{R}[\sigma^2(\mathbf{X}'\mathbf{X})^{-1}]\mathbf{R}'\}^{-1} \, (\mathbf{Rb} - \mathbf{q}) \sim \chi^2_J$

(2) Denominator: $(T-k) * s^2/\sigma^2 \sim \chi^2_{T-k}$

$$F = \frac{\chi^2_J / J}{\chi^2_T / T} = \frac{[(\mathbf{Rb} - \mathbf{q})' \, \{\mathbf{R}[\sigma^2(\mathbf{X}'\mathbf{X})^{-1}]\mathbf{R}'\}^{-1} \, (\mathbf{Rb} - \mathbf{q})\,]/J}{[(T-k)*s^2/\sigma^2]/(T-k)} \sim F_{J,T-k}.$$

Review: Testing H_0: $R\beta - q = 0$ with a Wald Test

Example: We test in the 3 FF factor model for IBM returns (T=569). Steps

1. H_0: $\beta_{SMB} = 0.2$ and $\beta_{HML} = 0.6$.
H_1: $\beta_{SMB} \neq 0.2$ and/or $\beta_{HML} \neq 0.6$. $\Rightarrow J = 2$

We define **R** (2x4) below and write $\mathbf{m} = \mathbf{R}\beta - \mathbf{q} = \mathbf{0}$:

$$\begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} * \begin{bmatrix} \beta_1 \\ \beta_{Mkt} \\ \beta_{SMB} \\ \beta_{HML} \end{bmatrix} = \begin{bmatrix} 0.2 \\ 0.6 \end{bmatrix}$$

2. Test-statistic: $F = W^*/J = (\mathbf{Rb} - \mathbf{q})' \{\mathbf{R}[s^2(\mathbf{X'X})^{-1}]\mathbf{R'}\}^{-1} (\mathbf{Rb} - \mathbf{q})$

Distribution under H_0: $F = W^*/2 \sim F_{2,T-4}$

(or asymptotic, $2*F \xrightarrow{d} \chi^2_2$)

Review: Testing H_0: $R\beta - q = 0$ with a Wald Test

Example (continuation): We use the R package *car* to test H_0.

```
library(car)
linearHypothesis(fit_ibm_ff3, c("SMB = 0.2","HML = 0.6"), test="F") # "F": exact test

Linear hypothesis test

Hypothesis:
SMB = 0.2
HML = 0.6

Model 1: restricted model
Model 2: ibm_x ~ Mkt_RF + SMB + HML

  Res.Df    RSS Df Sum of Sq      F    Pr(>F)
1    567 2.2691
2    565 1.9324  2   0.33667 49.217 < 2.2e-16 ***
```

⇒ reject H_0 at 5% level

Review: Testing H_0: $R\beta - q = 0$ with a Wald Test

Example (continuation): The asymptotic test uses test="Chisq".

```
> linearHypothesis(fit_ibm_ff3, c("SMB = 0.2","HML = 0.6"), test="Chisq")     # Asymptotic F
Linear hypothesis test

Hypothesis:
SMB = 0.2
HML = 0.6

Model 1: restricted model
Model 2: ibm_x ~ Mkt_RF + SMB + HML

  Res.Df    RSS Df Sum of Sq  Chisq Pr(>Chisq)
1    567 2.2691
2    565 1.9324  2   0.33667 98.433  < 2.2e-16 ***          ⇒ reject H0 at 5% level
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

qf(.95, df1=J, df2=(T - k))              # asymptotic distribution (Chi-square-distribution)
[1] 5.991465                              F_t_asym > 5.991465⇒ reject H0 at 5% level
```

The F Test: H_0: $R\beta - q = 0$

Approach(2): We know that imposing the restrictions leads to a loss of fit. R^2 must go down. Does it go down a lot? –i.e., significantly?

Recall (i) $\mathbf{e}^* = (\boldsymbol{y} - \mathbf{X}\mathbf{b}^*) = \boldsymbol{y} + (\mathbf{X}\mathbf{b} - \mathbf{X}\mathbf{b}) - \mathbf{X}\mathbf{b}^* = \mathbf{e} - \mathbf{X}(\mathbf{b}^* - \mathbf{b})$

(ii) $\mathbf{b}^* = \mathbf{b} - (\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'[\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}']^{-1}(\mathbf{R}\mathbf{b} - \mathbf{q})$

$$\Rightarrow \quad \mathbf{e}^{*\prime}\mathbf{e}^* = \mathbf{e}'\mathbf{e} + (\mathbf{b}^* - \mathbf{b})'\,\mathbf{X}'\mathbf{X}\,(\mathbf{b}^* - \mathbf{b})$$

Replacing $(\mathbf{b}^* - \mathbf{b})$ from (ii) in the above formula, we get:

$$\mathbf{e}^{*\prime}\mathbf{e}^* - \mathbf{e}'\mathbf{e} = (\mathbf{R}\mathbf{b} - \mathbf{q})'[\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}']^{-1}(\mathbf{R}\mathbf{b} - \mathbf{q})$$

Note: $\mathbf{e}^{*\prime}\mathbf{e}^* - \mathbf{e}'\mathbf{e}$ is a quadratic form, then we can use a lot of results for quadratic forms to derive its asymptotic distribution.

- Recall, the F-distribution is a ratio of two independent χ^2_J and χ^2_T RV divided by their degrees of freedom: $F = \dfrac{\chi^2_J / J}{\chi^2_T / T} \sim F_{J,T}$

The F Test: H_0: $R\beta - q = 0$

Then, to get to the F-test, we rely on two results:

– $W = (\mathbf{Rb} - \mathbf{q})'\{\mathbf{R}[\sigma^2(\mathbf{X}'\mathbf{X})^{-1}]\mathbf{R}'\}^{-1}(\mathbf{Rb} - \mathbf{q}) \sim \chi^2_J$ (if σ^2 is known)

– $\mathbf{e}'\mathbf{e}/\ \sigma^2 \sim \chi^2_{T-k}$.

$$\Rightarrow F = \frac{(\mathbf{e}^{*\prime}\mathbf{e}^{*} - \mathbf{e}'\mathbf{e})/J}{[\mathbf{e}'\mathbf{e}/(T-k)]} \sim F_{J,T-k}.$$

• We can write the F-test in terms of R^2's. Let

R^2 = unrestricted model = 1 – RSS/TSS

R^{*2} = restricted model fit = 1 – RSS*/TSS

Then, dividing and multiplying F by TSS we get:

$$F = \frac{(1 - R^{*2}) - (1-R^2)/J}{(1-R^2)/(T-k)} \sim F_{J,T-k}$$

or

$$F = \frac{(R^2 - R^{*2})/J}{(1-R^2)/(T-k)} \sim F_{J,T-k}.$$

The F Test: H_0: F-test of Goodness of Fit

• In the linear model, with a constant ($\mathbf{X}_1 = \boldsymbol{i}$):

$$\mathbf{y} = \mathbf{X}\,\beta + \varepsilon = \beta_1 + \mathbf{X}_2\,\beta_2 + \mathbf{X}_3\,\beta_3 + \ldots + \mathbf{X}_k\,\beta_k + \varepsilon$$

• We want to test if the slopes of $\mathbf{X}_2, \ldots, \mathbf{X}_k$ are equal to zero. That is,

H_0: $\beta_2 = \cdots = \beta_k = 0$

H_1: at least one $\beta \neq 0$ $\qquad \Rightarrow J = k - 1$

• We can write H_0: $\mathbf{R}\beta - \mathbf{q} = \mathbf{0}$ $\quad \Rightarrow \begin{bmatrix} 0 & 1 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \ldots \\ \beta_k \end{bmatrix} = \begin{bmatrix} 0 \\ \ldots \\ 0 \end{bmatrix}$

• We have $J = k - 1$. Then,

$$F = \{\, (R^2 - R^{*2})/(k-1) \,\} \,/\, [(1 - R^2)/(T-k)] \sim F_{k-1,T-k}.$$

• For the restricted model, $R^{*2} = 0$.

10

The F Test: H_0: F-test of Goodness of Fit

Then, $F = \frac{R^2/(k-1)}{(1-R^2)/(T-k)} \sim F_{k-1,T-k}$.

- Recall ESS/TSS = R^2 & RSS/TSS = (1 – R^2), we compute F:

$$F = \frac{R^2/(k-1)}{(1-R^2)/(T-k)} = \frac{\frac{ESS}{TSS}/(k-1)}{\frac{RSS}{TSS}/(T-k)}$$

$$F = \frac{ESS/(k-1)}{RSS/(T-k)}$$

- This test statistic is called the *F-test of goodness of fit*. It is reported in all regression packages as part of the regression output. In R, the lm function reports it as "*F-statistic*."

10

The F Test: H_0: F-test of Goodness of Fit

Example: We want to test if all the FF factors (Market, SMB, HML) are significant (J=3), using monthly data 1973 – 2020 (T=569).

```
y <- ibm_x
T <- length(ibm_x)
x0 <- matrix(1,T,1)
x <- cbind(x0, Mkt_RF,  SMB, HML)
k <- ncol(x)
b <- solve(t(x)%*% x)%*% t(x)%*%y                    # OLS regression
e <- y - x%*%b
RSS <- as.numeric(t(e)%*%e)
R2 <- 1 - as.numeric(RSS)/as.numeric(t(y)%*%y)    #R-squared
> R2
[1] 0.338985
> F_goodfit <- (R2/(k-1))/((1-R2)/(T-k))             # F-test of goodness of fit.
> F_goodfit
[1] 96.58204
```

⇒ F_goodfit > $F_{3,565,.05}$ = **2.62068** ⇒ Reject H_0.

10

The F Test: General Case – Example

- In the linear model

$$\mathbf{y} = \mathbf{X}\,\boldsymbol{\beta} + \boldsymbol{\varepsilon} = \beta_1 + \mathbf{X}_2\,\beta_2 + \mathbf{X}_3\,\beta_3 + \mathbf{X}_4\,\beta_4 + \boldsymbol{\varepsilon}$$

- We want to test if the slopes $\mathbf{X}_3, \mathbf{X}_4$ are equal to zero. That is,

H_0: $\beta_3 = \beta_4 = \mathbf{0}$

H_1: $\beta_3 \neq \mathbf{0}$ or $\beta_4 \neq \mathbf{0}$ or both β_3 and $\beta_4 \neq \mathbf{0}$

- We can use, $F = (\mathbf{e}^{*\prime}\mathbf{e}^* - \mathbf{e}'\mathbf{e})/J \;/\; [\mathbf{e}'\mathbf{e}/(T-k)] \sim F_{J,T-k}$.

- Define $\mathbf{y} = \mathbf{X}\,\boldsymbol{\beta} + \boldsymbol{\varepsilon} = \beta_1 + \mathbf{X}_2\,\beta_2 + \boldsymbol{\varepsilon}$ (RSS_R)

$\mathbf{y} = \beta_1 + \mathbf{X}_2\,\beta_2 + \mathbf{X}_3\,\beta_3 + \mathbf{X}_4\,\beta_4 + \boldsymbol{\varepsilon}$ (RSS_U)

$$F(k_U - k_R, T-k) = \frac{\dfrac{RSS_R - RSS_U}{(k_U - k_R)}}{\dfrac{RSS_U}{(T-k_U)}}$$

32

The F Test: Are SMB and HML Priced Factors?

Example: We want to test if the additional FF factors (SMB, HML) are significant, using monthly data 1973 – 2020 (T=569).

Unrestricted Model:

(U) $(r_{IBM,t} - r_f) = \beta_0 + \beta_1\,(r_{m,t} - r_f) + \beta_2\,\text{SMB} + \beta_3\,\text{HML} + \varepsilon$

Hypothesis: H_0: $\beta_2 = \beta_3 = 0$

H_1: $\beta_2 \neq 0$ and/or $\beta_3 \neq 0$

Then, the Restricted Model:

(R) $(r_{IBM,t} - r_f) = \beta_0 + \beta_1\,(r_{m,t} - r_f) + \varepsilon$

Test: $F = \dfrac{(RSS_R - RSS_U)/J}{RSS_U/(T-k_u)} \sim F_{J,T-k}$. with $J = k_U - k_R = 4 - 2 = 2$

The F Test: Are SMB and HML Priced Factors?

Example (continuation): The unrestricted model was already estimated in Lecture 3. For the restricted model:

```
y <- ibm_x
x0 <- matrix(1,T,1)
x_r <- cbind(x0,Mkt_RF)                          # Restricted X vector
k <- ncol(x)
T <- nrow(x)
k2 <- ncol(x_r)

b2 <- solve(t(x_r)%*% x_r)%*% t(x_r)%*%y        # Restricted OLS regression
e2 <- y – x_r%*%b2
RSS2 <- as.numeric(t(e2)%*%e2)
> RSS = 1.932442                                 # RSS_U
> RSS2 = 1.964844                                # RSS_R
> J <- k - k2                                    # J = degrees of freedom of numerator
> F_test <- ((RSS2 - RSS)/J)/(RSS/(T-k))
```

The F Test: Are SMB and HML Priced Factors?

Example (continuation):

```
> F_test <- ((RSS2 - RSS)/J)/(RSS/(T-k))
> F_test
[1] 4.736834
> qf(.95, df1=J, df2=(T-k))                      # F_{2,565,.05} value (≈ 3)
[1] 3.011672                                     ⇒ Reject H0.
> p_val <- 1 - pf(F_test, df1=J, df2=(T-k))      # p-value of F_test
> p_val
[1] 0.009117494                                  ⇒ p-value is small ⇒ Reject H0.
```


The F Test: Are SMB and HML Priced Factors?

Example (continuation):

There is package in R, *lmtest,* that performs this test, *waldtest,* (and many others, used in this class). You need to install it first.

Note: The models need to be nested. For the *waldtest*, the default reports the *F-test* with the F distribution.

```
library(lmtest)
fit_wU <- lm (ibm_x ~ Mkt_RF + SMB + HML)
fit_wR <- lm (ibm_x ~ Mkt_RF)
waldtest(fit_wU, fit_wR)

Wald test

Model 1: ibm_x ~ Mkt_RF + SMB + HML
Model 2: ibm_x ~ Mkt_RF
  Res.Df Df      F   Pr(>F)
1    565
2    567 -2 4.7368 0.009117 **
```

⇒ p-value is small: Reject H_0

F-test: Two Categories & The Chow Test

• Suppose we are interested in the effect of gender on CEO's compensation. We have data on CEO's compensation (y) and CEO's gender, along with CEO's experience (X_1), sales of the CEO's company (X_2), and profitability (X_3).

• We hypothesize that gender matter. Then, we estimate two models, one for each gender:

M1 $\quad y_i = \beta_0^1 + \beta_1^1 X_{1,I} + \beta_2^1 X_{2,i} + \beta_3^1 X_{3,i} + \varepsilon_i$ for i = *Male*

M2 $\quad y_i = \beta_0^2 + \beta_1^2 X_{1,I} + \beta_2^2 X_{2,i} + \beta_3^2 X_{3,i} + \varepsilon_i$ for $i \neq$ *Female*

• Alternatively, we estimate only one model ("*pooling*"). That is, gender does not affect a CEO's compensation. Then, we estimate:

Pooled $\quad y_i = \beta_0 + \beta_1 X_{1,i} + \beta_2 X_{2,i} + \beta_3 X_{3,i} + \varepsilon_i$ for all i

Q: Which model should we use?

26

F-test: Two Categories & The Chow Test

• We test H_0 (No *gender differences*): $\beta_0^1 = \beta_0^2 = \beta_0$
$\beta_1^1 = \beta_1^2 = \beta_1$
$\beta_2^1 = \beta_2^2 = \beta_2$
$\beta_3^1 = \beta_3^2 = \beta_3$

H_1 (*gender differences*): For at least k (= 0, 1, 2, 3): $\beta_k^1 \neq \beta_k^2$

• An F-Test can be used to test H_0:

- The pooled estimation is the Restricted estimation
- The two estimations (by gender) are the Unrestricted estimation.

• The F-test constructed using a variable that can divide the data into 2 categories to compute RSS_R & RSS_U is usually referred as *Chow test*.

27

F-test: Two Categories & The Chow Test

• A Chow Test is used to test if a variable that can divide the data into 2 categories matters. That is, a Chow test checks if we need only one model ("*pooling*") for both categories or not.

• Chow Test (an F-test) –Chow (1960, *Econometrica*):

(1) Run OLS with all the data, with no distinction between categories. (Pooled regression or Restricted regression). Keep RSS_R.

(2) Run two separate OLS, one for each category (Unrestricted regression). Keep RSS_1 and RSS_2 $\Rightarrow RSS_U = RSS_1 + RSS_2$.

(3) Run a standard F-test (testing Restricted vs. Unrestricted models):

$$F = \frac{(RSS_R - RSS_U)/(k_U - k_R)}{(RSS_U)/(T - k_U)} = \frac{(RSS_R - [RSS_1 + RSS_2])/k}{(RSS_1 + RSS_2)/(T - 2k)}$$

28

Chow Test: Males or Females visit doctors more?

German Health Care Usage Data, 7,293 Individuals, Varying Numbers of Periods
Variables in the file are
Data downloaded from Journal of Applied Econometrics Archive. This is an unbalanced panel with 7,293 individuals. There are altogether 27,326 observations. The number of observations ranges from 1 to 7 per family. (Frequencies are: 1=1525, 2=2158, 3=825, 4=926, 5=1051, 6=1000, 7=987). The dependent variable of interest is

DOCVIS = number of visits to the doctor in the observation period

HHNINC = household nominal monthly net income in German marks / 10000.
(4 observations with income=0 were dropped)
HHKIDS = children under age 16 in the household = 1; otherwise = 0
EDUC = years of schooling
AGE = age in years
MARRIED= marital status (1 = if married)
WHITEC = 1 if has "white collar" job

29

Chow Test: Males or Females visit doctors more?

- OLS Estimation for **Men** only. Keep RSS_M = **379.8470**

```
+----------------------------------------------------+
| Ordinary    least squares regression               |
| LHS=HHNINC   Mean                 =   .3590541     |
|              Standard deviation   =   .1735639     |
|              Number of observs.   =      14243     |
| Model size   Parameters           =          5     |
|              Degrees of freedom   =      14238     |
| Residuals    Sum of squares       =   379.8470     |
|              Standard error of e  =   .1633352     |
| Fit          R-squared            =   .1146423     |
|              Adjusted R-squared   =   .1143936     |
+----------------------------------------------------+
```

Variable	Coefficient	Standard Error	b/St.Er.	P[|Z|>z]	Mean of X
Constant	.04169***	.00894	4.662	.0000	
AGE	.00086***	.00013	6.654	.0000	42.6528
EDUC	.02044***	.00058	35.528	.0000	11.7287
MARRIED	.03825***	.00341	11.203	.0000	.76515
WHITEC	.03969***	.00305	13.002	.0000	.29994

30

Chow Test: Males or Females visit doctors more?

- OLS Estimation for **Women** only. Keep RSS_W = **363.8789**

```
+----------------------------------------------------+
| Ordinary    least squares regression               |
| LHS=HHNINC   Mean                 =    .3444951    |
|              Standard deviation   =    .1801790    |
|              Number of observs.   =       13083    |
| Model size   Parameters           =           5    |
|              Degrees of freedom   =       13078    |
| Residuals    Sum of squares       =    363.8789    |
|              Standard error of e  =    .1668045    |
| Fit          R-squared            =    .1432098    |
|              Adjusted R-squared   =    .1429477    |
+----------------------------------------------------+
+--------+--------------+----------------+--------+--------+----------+
|Variable| Coefficient  | Standard Error |b/St.Er.|P[|Z|>z]| Mean of X|
+--------+--------------+----------------+--------+--------+----------+
|Constant|     .01191           .01158        1.029   .3036          |
|AGE     |     .00026*          .00014        1.875   .0608   44.4760|
|EDUC    |     .01941***        .00072       26.803   .0000   10.8764|
|MARRIED |     .12081***        .00343       35.227   .0000    .75151|
|WHITEC  |     .06445***        .00334       19.310   .0000    .29924|
+--------+------------------------------------------------------------+
```

31

Chow Test: Males or Females visit doctors more?

```
+----------------------------------------------------+
| Ordinary    least squares regression               |
| LHS=HHNINC   Mean                 =    .3520836    |
|              Standard deviation   =    .1769083    |
|              Number of observs.   =       27326    |
| Model size   Parameters           =           5    |
|              Degrees of freedom   =       27321    |
| Residuals    Sum of squares       =    752.4767    | All
| Residuals    Sum of squares       =    379.8470    | Men
| Residuals    Sum of squares       =    363.8789    | Women
+----------------------------------------------------+
+--------+--------------+----------------+--------+--------+----------+
|Variable| Coefficient  | Standard Error |b/St.Er.|P[|Z|>z]| Mean of X|
+--------+--------------+----------------+--------+--------+----------+
|Constant|     .04186***        .00704        5.949   .0000          |
|AGE     |     .00030***     .919581D-04      3.209   .0013   43.5257|
|EDUC    |     .01967***        .00045       44.180   .0000   11.3206|
|MARRIED |     .07947***        .00239       33.192   .0000    .75862|
|WHITEC  |     .04819***        .00225       21.465   .0000    .29960|
+--------+------------------------------------------------------------+
```

Chow Test = $F = [(752.4767 - (379.847 + 363.8789))/5] \,/\, [(379.847 + 363.8789)/(27{,}326 - 10)] = 64.281$

32

$F(5, 27311) = 2.214100$ $\Rightarrow$ reject H_0

F-Test: Structural Change & Chow Test

• Suppose there is an event that we think had a big effect on the behaviour of our model. Suppose the event occurred at time T_{SB}. We think that the before and after behaviour of the model is significantly different. For example, the parameters are different before and after T_{SB}. That is,

$$y_i = \beta_0^1 + \beta_1^1 X_{1,I} + \beta_2^1 X_{2,i} + \beta_3^1 X_{3,i} + \varepsilon_i \qquad \text{for } i \leq T_{SB}$$
$$y_i = \beta_0^2 + \beta_1^2 X_{1,I} + \beta_2^2 X_{2,i} + \beta_3^2 X_{3,i} + \varepsilon_i \qquad \text{for } i > T_{SB}$$

The event caused *structural change* in the model. T_{SB} separates the behaviour of the model in two regimes/categories ("*before*" & "*after*".)

• A Chow test tests if one model applies to both regimes:

$$y_i = \beta_0 + \beta_1 X_{1,i} + \beta_2 X_{2,i} + \beta_3 X_{3,i} + \varepsilon_i \qquad \text{for all } i$$

• Under H_0 (No *structural change*), the parameters are the same for all i.

33

F-Test: Structural Change & Chow Test

• We test H_0 (No *structural change*):

$$\beta_0^1 = \beta_0^2 = \beta_0$$
$$\beta_1^1 = \beta_1^2 = \beta_1$$
$$\beta_2^1 = \beta_2^2 = \beta_2$$
$$\beta_3^1 = \beta_3^2 = \beta_3$$

H_1 (*structural change*): For at least k (= 0, 1, 2, 3): $\beta_k^1 \neq \beta_k^2$

• What events may have this effect on a model? A financial crisis, a big recession, an oil shock, Covid-19, etc.

• Testing for structural change is the more popular use of Chow tests.

• Chow tests have many interpretations: tests for structural breaks, pooling groups, parameter stability, predictive power, etc.

• One important consideration: T may not be large enough.

34

F-Test: Structural Change & Chow Test

- We structure the Chow test to test H_0 (No *structural change*), as usual.

- Steps for Chow (Structural Change) Test:

(1) Run OLS with all the data, with no distinction between regimes. (Restricted or pooled model). Keep RSS_R.

(2) Run two separate OLS, one for each regime (Unrestricted model):

Before Date T_{SB}. Keep RSS_1.

After Date T_{SB}. Keep RSS_2. $\Rightarrow RSS_U = RSS_1 + RSS_2$.

(3) Run a standard F-test (testing Restricted vs. Unrestricted models):

$$F = \frac{(RSS_R - RSS_U)/(k_U - k_R)}{(RSS_U)/(T - k_U)} = \frac{(RSS_R - [RSS_1 + RSS_2])/k}{(RSS_1 + RSS_2)/(T - 2k)}$$

35

F-Test: Structural Change & Chow Test

Example: We test if the Oct 1973 oil shock in quarterly GDP growth rates had an structural change on the GDP growth rate model.

We model GDP the growth rate with an AR(1) model, that is, GDP growth rate depends only on its own lagged growth rate:

$$y_t = \beta_0 + \beta_1 \, y_{t-1} + \varepsilon_t$$

```
GDP_da <- read.csv("http://www.bauer.uh.edu/rsusmel/4397/GDP_q.csv", head=TRUE,
sep=",")
x_date <- GDP_da$DATE
x_gdp <- GDP_da$GDP
x_dummy <- GDP_da$D73
T <- length(x_gdp)
t_s <- 108                          # T_SB = Oct 1973

lr_gdp <- log(x_gdp[-1]/x_gdp[-T])
T <- length(lr_gdp)
lr_gdp0 <- lr_gdp[-1]
lr_gdp1 <- lr_gdp[-T]
t_s <- t_s -1                       # Adjust t_s (we lost the first observation)
```

36

F-Test: Structural Change & Chow Test

Example (continuation):

```
y <- lr_gdp0
x1 <-  lr_gdp1
T <- length(y)
x0 <- matrix(1,T,1)
x <- cbind(x0,x1)
k <- ncol(x)

# Restricted Model (Pooling all data)
fit_ar1 <- lm(lr_gdp0 ~ lr_gdp1)                 # Fitting AR(1) (Restricted) Model
e_R <- fit_ar1 $residuals                        # regression residuals, e
RSS_R <- sum(e_R^2)                              # RSS Restricted

> summary(fit_ar1)

Coefficients:
            Estimate   Std. Error t value Pr(>|t|)
(Intercept) 0.011565   0.001145  10.097  < 2e-16 ***
lr_gdp1     0.244846   0.056687   4.319 2.14e-05 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.01296 on 294 degrees of  freedom
```

37

F-Test: Structural Change & Chow Test

Example (continuation):

```
# Unrestricted Model (Two regimes)

y_1 <- y[1:t_s]
x_u1 <- x[1:t_s,]
fit_ar1_1 <- lm(y_1 ~ x_u1 - 1)                 # AR(1) Regime 1
e1 <- fit_ar1_1$residuals                       # Regime 1 regression residuals, e
RSS1 <- sum(e1^2)                               # RSS Regime 1

kk = t_s+1                                      # Starting date for Regime 2
y_2 <- y[kk:T]
x_u2 <- x[kk:T,]
fit_ar1_2 <- lm(y_2 ~ x_u2 - 1)                 # AR(1) Regime 2
e2 <- fit_ar1_2$residuals                       # Regime 2 regression residuals, e
RSS2 <- sum(e2^2)                               #  RSS Regime 2

F <- ((RSS_R - (RSS1+RSS2))/k)/((RSS1+RSS2)/(T - 2*k))
> F
[1] 4.877371
p_val <- 1 - pf(F, df1 = 2, df2 = T - 2*k)  # p-value of F_test
> p_val
[1] 0.00824892
```

⇒ small p-values: Reject H_0 (No *structural change*).

38

3

F-Test: Structural Change & Chow Test

Example: 3 Factor Fama-French Model for IBM (continuation)

Q: Did the dot.com bubble (end of 2001) affect the structure of the FF Model? Sample: Jan 1973 – June 2020 (T = 569).

Pooled RSS = **1.9324**

Jan 1973 – Dec 2001 RSS = RSS_1 = **1.33068** (T = 342)

Jan 2002 – June 2020 RSS = RSS_2 = **0.57912** (T = 227)

$$F = \frac{[RSS_R - (RSS_1 + RSS_2)]/k}{(RSS_1 + RSS_2)/(T-k)} = \frac{[1.9324 - (1.3307 + 0.57911)]/4}{(1.3307 + 0.57911)/(569 - 2*4)} = 1.6627$$

$\Rightarrow$ Since $F_{4,565,.05} = 2.39$, we cannot reject H_0

	Constant	Mkt – rf	SMB	HML	RSS	T
1973-2020	-0.0051	0.9083	-0.2125	-0.1715	**1.9324**	569
1973-2001	-0.0038	0.8092	-0.2230	-0.1970	**1.3307**	342
2002 – 2020	-0.0073	1.0874	-0.1955	-0.3329	**0.5791**	227

39

Testing Model Specification: Nested Models

• In previous examples, we have two nested models, one is the restricted version of the other (Fama-French 3-factor model vs CAPM). In the case of omitted variables:

(U) $\quad \mathbf{y} = \mathbf{X}\,\boldsymbol{\beta}_1 + \mathbf{Z}\,\boldsymbol{\beta}_2 + \boldsymbol{\varepsilon}$ $\quad$ –the "long regression,"

(R) $\quad \mathbf{y} = \mathbf{X}\,\boldsymbol{\beta}_1 + \boldsymbol{\varepsilon}$ $\quad$ –the "short regression."

To test H_0 (No omitted variables): $\boldsymbol{\beta}_2 = 0$, we can use the F-test:

$$F = \frac{(RSS_R - RSS_U)/J}{RSS_U/(T-k)} \sim F_{J,T-K}.$$

Example: We have performed this F-test to test if in the 3-factor FF model for IBM returns, **SMB** and **HML** were significant, which they were. That is, we showed that the usual CAPM formulation for IBM returns had omitted variables: **SMB** & **HML**.

40

Testing Model Specification: Non-Nested Models

- So far, all our tests (t-, F- & Wald tests) have been based on nested models, where the R model is a restricted version of the U model.

Example:

Model U $\quad \mathbf{Y} = \mathbf{X}\beta + \mathbf{W}\delta + \varepsilon$ (Unrestricted)

Model R $\quad \mathbf{Y} = \mathbf{X}\beta + \xi$ (Restricted)

Model U becomes Model R under H_0: $\boldsymbol{\delta} = \mathbf{0}$.

- Sometimes, we have two rival models to choose between, where neither can be nested within the other -i.e., neither is a restricted version of the other.

Example:

Model 1 $\quad \mathbf{Y} = \mathbf{X}\beta + \mathbf{W}\delta + \varepsilon$

Model 2 $\quad \mathbf{Y} = \mathbf{X}\beta + \mathbf{Z}\gamma + \xi$

Testing Model Specification: Non-Nested Models

Example:

Model 1 $\quad \mathbf{Y} = \mathbf{X}\beta + \mathbf{W}\delta + \varepsilon$

Model 2 $\quad \mathbf{Y} = \mathbf{X}\beta + \mathbf{Z}\gamma + \xi$

- If the dependent variable is the same in both models (as is the case here), we can simply use Adjusted-R^2 to rank the models and select the one with the largest Adjusted-R^2.

- We can also use AIC and/or BIC to rank the models.

- But, we can also use more sophisticated, testing-based, methods.

Non-nested Models and Tests: Encompassing

• **Testing-based Method 1:** Encompassing

(1) Form a composite or *encompassing* model that nests both rival models −Model 1 & Model 2. This is the **unrestricted Model**, **ME**.

(2) Test the relevant restrictions of each rival model against **ME**. We do two F-tests:

(i) Test **ME** (Unrestricted Model) against Model 1 (Restricted Model)
(ii) Test **ME** (Unrestricted Model) against Model 2 (Restricted Model)

• If we reject the restrictions against one Model, say Model 1, and we cannot reject the restrictions against the other, Model 2, we are done: We select the Model that the F test do not reject restrictions (Model 2).

Assuming the restrictions cannot be rejected, we prefer the model with the lower F statistic for the test of restrictions.

Non-nested Models and Tests: Encompassing

Example: We have:

Model 1 $\mathbf{Y} = \mathbf{X}\beta + \mathbf{W}\delta + \varepsilon$
Model 2 $\mathbf{Y} = \mathbf{X}\beta + \mathbf{Z}\gamma + \xi$

Then, the **Encompassing Model** (**ME**) is:

ME: $\mathbf{Y} = \mathbf{X}\beta + \mathbf{W}\delta + \mathbf{Z}\gamma + \varepsilon$

Now test, separately, the hypotheses (1) $\boldsymbol{\delta} = \mathbf{0}$ and (2) $\boldsymbol{\gamma} = \mathbf{0}$. That is,

F-test for H_0: $\boldsymbol{\gamma} = \mathbf{0}$: ME (U Model) vs Model 1 (R Model).
F-test for H_0: $\boldsymbol{\delta} = \mathbf{0}$: ME (U Model) vs Model 2 (R Model).

If we reject H_0: $\boldsymbol{\gamma} = \mathbf{0}$ ⇒ We have evidence against Model 1
If we reject H_0: $\boldsymbol{\delta} = \mathbf{0}$ ⇒ We have evidence against Model 2.

Note: We test a hybrid model, a combination of two models. Also, multicollinearity may appear.

Non-nested Models and Tests: IFE or PPP?

• Two of the main theories to explain the behaviour of exchange rates, S_t, are the **International Fisher Effect** (**IFE**) and the **Purchasing Power Parity** (**PPP**). We use the direct notation for S_t, that is, units of Domestic Currency per 1 unit of Foreign currency.

• IFE states that, in equilibrium, changes in exchange rates (**e**) are driven by the interest rates differential between the domestic currency, i_d, and the foreign currency, i_f:. A DGP consistent with IFE is:

$$\mathbf{e} = \alpha^1 + \beta^1 (\mathbf{i_d} - \mathbf{i_f}) + \varepsilon^1$$

• Relative PPP states that that, in equilibrium, **e** are driven by the inflation rates differential between the domestic Inflation rate, I_d, and the foreign Inflation rate, I_f. A GDP consistent with IFE is:

$$\mathbf{e} = \alpha^2 + \beta^2 (\mathbf{I_d} - \mathbf{I_f}) + \varepsilon^2$$

• Theories are non-nested, use non-nested methods to pick a model.

Non-nested Models and Tests: IFE or PPP?

Example: We apply hat drives log changes in exchange rates for the USD/GBP (e): $(i_d - i_f)$ or $(I_d - I_f)$?

Model 1 (IFE): $\mathbf{e} = \alpha^1 + \beta^1 (\mathbf{i_d} - \mathbf{i_f}) + \varepsilon^1$

Model 2 (PPP): $\mathbf{e} = \alpha^2 + \beta^2 (\mathbf{I_d} - \mathbf{I_f}) + \varepsilon^2$

```
SF_da <- read.csv("http://www.bauer.uh.edu/rsusmel/4397/SpFor_prices.csv", head=TRUE, sep=",")
x_date <- SF_da$Date
x_S <- SF_da$GBPSP
x_F3m <- SF_da$GBP3M
i_us3 <- SF_da$Dep_USD3M
i_uk3 <- SF_da$Dep_UKP3M
cpi_uk <- SF_da$UK_CPI
cpi_us <- SF_da$US_CPI
T <- length(x_S)
int_dif <- (i_us3[-1] - i_uk3[-1])/100
lr_usdgbp <- log(x_S[-1]/x_S[-T])
I_us <- log(cpi_us[-1]/cpi_us[-T])
I_uk <- log(cpi_uk[-1]/cpi_uk[-T])
inf_dif <- (I_us - I_uk)
```


Non-nested Models and Tests: IFE or PPP?

Example (continuation): Encompassing Model (U Model)

$$\mathbf{e} = \alpha + \beta_1 (\mathbf{i_d} - \mathbf{i_f}) + \beta_2 (\mathbf{I_d} - \mathbf{I_f}) + \varepsilon^1$$

```
# Encompassing Model and Test
fit_me <- lm(lr_usdgbp ~ int_dif + inf_dif)
> summary(fit_me)

Coefficients:
             Estimate    Std. Error  t value Pr(>|t|)
(Intercept)  -0.0009633  0.0016210   -0.594   0.5527
int_dif      -0.0278510  0.0741189   -0.376   0.7073
inf_dif       0.7444711  0.3429106    2.171   0.0306 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.02662 on 360 degrees of freedom
Multiple R-squared:  0.01316,   Adjusted R-squared:  0.007673
F-statistic: 2.399 on 2 and 360 DF,  p-value: 0.09221
```

int_dif: ⇒ cannot reject H_0: $\beta_1 = 0$.
inf_dif: ⇒ reject H_0: $\beta_2 = 0$.

Note: Two F-tests are needed, but for the one variable case, the t-tests are equivalent.

Non-nested Models and Tests: IFE or PPP?

Example (continuation): The package in R, *lmtest,* performs this test, *encomptest*. Recall you need to install it first: install.packages("lmtest").

Note: The test reported is an F-test ~ $F_{1,T-k}$, which, in this case with only one variable in each Model, is equal to $(t_{T-k})^2$.

```
library(lmtest)
fit_m1 <- lm(lr_usdgbp ~ int_dif)          # Restricted Model 1 (IFE)
fit_m2 <- lm(lr_usdgbp ~ inf_dif)          # Restricted Model 2 (PPP)
> encomptest(fit_m1, fit_m2)

1: lr_usdgbp ~ int_dif
Model 2: lr_usdgbp ~ inf_dif
Model E: lr_usdgbp ~ int_dif + inf_dif
          Res.Df Df      F       Pr(>F)
M1 vs. ME    360 -1 4.7134 0.03058 *
M2 vs. ME    360 -1 0.1412 0.70732
```

M1 vs. ME: ⇒ reject H_0: $\beta_2 = 0$. Check: $(2.171)^2 = 4.713$

Non-nested Models and Tests: *J*-test

• **Testing-based Method 1:** Davidson-MacKinnon (1981)'s *J*-test.
We start with two non-nested models. Say,

Model 1: $\mathbf{Y} = \mathbf{X}\beta + \varepsilon$

Model 2: $\mathbf{Y} = \mathbf{Z}\gamma + \xi$

Idea: If Model 2 is true, then the fitted values from the Model 1, when added to the 2nd equation, should be insignificant.

• Steps:

(1) Estimate **Model 1** $\Rightarrow$ obtain fitted values: **Xb**.

(2) Add **Xb** to the list of regressors in Model 2

$\Rightarrow \mathbf{Y} = \mathbf{Z}\gamma + \lambda\mathbf{Xb} + \xi$

(3) Do a *t-test* on λ. A significant *t*-value would be evidence against Model 2 and in favour of **Model 1**.

Non-nested Models and Tests: *J*-test

(4) Repeat the procedure for the models the other way round.

(4.1) Estimate **Model 2** $\Rightarrow$ obtain fitted values: **Zc**.

(4.2) Add **Zc** to the list of regressors in Model 1:

$\Rightarrow \mathbf{Y} = \mathbf{X}\beta + \lambda\,\mathbf{Zc} + \varepsilon$

(4.3) Do a *t-test* on λ. A significant *t*-value would be evidence against **Model 1** and in favour of **Model 2**.

(5) Rank the models on the basis of this test.

• It is possible that we cannot reject both models. This is possible in small samples, even if one model, say Model 2, is true.

• It is also possible that both *t-tests* reject H_0 ($\lambda \neq 0$ & $\lambda \neq 0$). This is not unusual. McAleer's (1995), in a survey, reports that out of 120 applications all models were rejected 43 times.

Non-nested Models and Tests: *J*-test

• Situations:

(1) Both OK: $\lambda = 0$ and $\lambda = 0$ ⇒ get more data

(2) Only 1 is OK: $\lambda \neq 0$ and $\lambda = 0$ (**Model 2** is OK)

$\lambda \neq 0$ and $\lambda = 0$ (**Model 1** is OK)

(3) Both rejected: $\lambda \neq 0$ and $\lambda \neq 0$ ⇒ new model is needed.

Technical Note: As some of the regressors in step (3) are stochastic, Davidson and MacKinnon (1981) show that the *t-test* is *asymptotically* valid.

Non-nested Models: *J*-test – IFE or PPP?

Example: Now, we test Model 1 vs Model 2, using the *J*-test.

Model 1 (IFE): $\mathbf{e} = \alpha^1 + \beta^1 (\mathbf{i}_d - \mathbf{i}_f) + \varepsilon^1$

Model 2 (PPP): $\mathbf{e} = \alpha^2 + \beta^2 (\mathbf{I}_d - \mathbf{I}_f) + \varepsilon^2$

```
y <- lr_usdgbp
fit_m1 <- lm( y ~ int_dif)
summary(fit_m1)
y_hat1  <- fitted(fit_m1)
fit_J1 <- lm( y ~ inf_dif + y_hat1)
summary(fit_J1)

fit_m2 <- lm( y ~ inf_dif)
summary(fit_m2)
y_hat2  <- fitted(fit_m2)
fit_J2 <- lm( y ~ int_dif + y_hat2)
summary(fit_J2)
```

Non-nested Models: *J*-test – IFE or PPP?

Example (continuation):

```
> fit_m1 <- lm( y ~ int_dif)
> y_hat1  <- fitted(fit_m1)
> fit_J1 <- lm(formula = y ~ inf_dif + y_hat1)
> summary(fit_J1)

Residuals:
     Min       1Q    Median       3Q       Max
-0.136310 -0.014168  0.000351  0.017227  0.092421

Coefficients:
              Estimate   Std. Error t value Pr(>|t|)
(Intercept)  0.0001497  0.0025556   0.059   0.9533
inf_dif      0.7444711  0.3429106   2.171   0.0306 *
y_hat1       1.2853298  3.4206106   0.376   0.7073
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.02662 on 360 degrees of freedom
Multiple R-squared:  0.01316,   Adjusted R-squared:  0.007673
F-statistic: 2.399 on 2 and 360 DF,  p-value: 0.09221
```

⇒ cannot reject H_0: λ=0. (Good for Model 2)

Non-nested Models: *J*-test – IFE or PPP?

Example (continuation):

```
> fit_m2 <- lm( y ~ inf_dif)
> y_hat2  <- fitted(fit_m2)
> fit_J2 <- lm(formula = y ~ int_dif + y_hat2)
> summary(fit_J2)

Residuals:
     Min       1Q    Median       3Q       Max
-0.136310 -0.014168  0.000351  0.017227  0.092421

Coefficients:
             Estimate    Std. Error t value Pr(>|t|)
(Intercept) -0.0003045  0.0016409  -0.186   0.8529
int_dif     -0.0278510  0.0741189  -0.376   0.7073
y_hat2       1.0066945  0.4636932   2.171   0.0306 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.02662 on 360 degrees of freedom
Multiple R-squared:  0.01316,   Adjusted R-squared:  0.007673
F-statistic: 2.399 on 2 and 360 DF,  p-value: 0.09221
```

⇒ Reject H_0: λ=0. (Again, good for Model 2)

Non-nested Models: *J*-test – IFE or PPP?

Example (continuation):
The *lmtest* package also performs this test. Recall that you need to install it first: install.packages("lmtest").

```
library(lmtest)
fit_m1 <- lm(lr_usdgbp ~ int_dif)
fit_m2 <- lm(lr_usdgbp ~ inf_dif)
> jtest(fit_m1, fit_m2)
J test

Model 1: lr_usdgbp ~ int_dif
Model 2: lr_usdgbp ~ inf_dif
                 Estimate Std. Error t value Pr(>|t|)
M1 + fitted(M2)  1.0067   0.4637   2.1710  0.03058 *
M2 + fitted(M1)  1.2853   3.4206   0.3758  0.70732
```

⇒ Reject H_0: λ=0. (Model 2 selected)

Non-nested Models: *J*-test – Remarks

• The *J*-test was designed to test non-nested models (one model is the true model, the other is the false model), not for choosing competing models –the usual use of the test.

• The *J*-test is likely to *over reject* the true (model) hypothesis when one or more of the following features is present:
i) A poor fit of the true model.
ii) A low/moderate correlation between the regressors of the 2 models.
iii) The false model includes more regressors than the correct model.

Davidson and MacKinnon (2004) state that the *J*-test will over-reject, *often quite severely* in finite samples when the sample size is small or where conditions (i) or (iii) above are obtained.

Testing Remarks: A word about α

- Ronald Fisher, before computers, tabulated distributions. He used a .10, .05, and .01 percentiles. These tables were easy to use and, thus, those percentile became the de-facto standard α for testing H_0.

- "It is usual and convenient for experimenters to take 5% as a standard level of significance." –Fisher (1934).

- Given that computers are powerful and common, why is $p = 0.051$ unacceptable, but $p = 0.049$ is great? There is no published work that provides a theoretical basis for the standard thresholds.

- Rosnow and Rosenthal (1989): " ... surely God loves .06 nearly as much as .05."

Testing Remarks: A word about α

<u>Practical advise</u>: In the usual Fisher's null hypothesis (significance) testing, significance levels, α, are arbitrary. Make sure you pick one, say 5%, and stick to it throughout your analysis or paper.

- Report *p-values*, along with CI's. Search for *economic significance*.

Testing Remarks: A word about H_0

• In applied work, we only learn when we reject H_0; say, when the *p-value* < α. But, rejections are of two types:

- Correct ones, driven by the power of the test.
- Incorrect ones, driven by Type I Error ("*statistical accident*," luck).

• It is important to realize that, however small the *p-value*, there is always a finite chance that the result is a pure accident. At the 5% level, there is 1 in 20 chances that the rejection of H_0 is just luck.

• Since negative results are difficult to publish (*publication bias*), there is an unknown but possibly large number of false claims taken as truths.

Example: If $\alpha = 0.05$, proportion of false H_0=10%, and $\pi = .50$, **47.4%** of rejections are true H_0 -i.e., "*false positives*."

Model Specification: Checking (A1)

60

OLS Estimation - Assumptions

• CLM Assumptions

(**A1**) DGP: $\mathbf{y} = \mathbf{X}\,\beta + \boldsymbol{\varepsilon}$ is correctly specified.

(**A2**) $E[\boldsymbol{\varepsilon}|\mathbf{X}] = 0$

(**A3**) $Var[\boldsymbol{\varepsilon}|\mathbf{X}] = \sigma^2\, \mathbf{I}_T$

(**A4**) $\mathbf{X}$ has full column rank –rank($\mathbf{X}$)=k-, where $T \geq k$.

Q: What happens when (**A1**) is not correctly specified?

• First, we look at (**A1**), in the context of linearity. Are we omitting a relevant regressor? Are we including an irrelevant variable? What happens when we impose restrictions in the DGP?

• Second, in (**A1**), we allow some non-linearities in its functional form.

Specification Errors: Omitted Variables

• Omitting relevant variables: Suppose the correct model (DGP) is

$\mathbf{y} = \mathbf{X}_1\boldsymbol{\beta}_1 + \mathbf{X}_2\boldsymbol{\beta}_2 + \boldsymbol{\varepsilon}$ –the "long regression," with $\mathbf{X}_1$ & $\mathbf{X}_2$.

But, we compute OLS omitting $\mathbf{X}_2$. That is,

$\mathbf{y} = \mathbf{X}_1\boldsymbol{\beta}_1 + \boldsymbol{\varepsilon}$ –the "short regression."

We have two *nested* models: one model becomes the other, once a restriction is imposed. In the above case, the true model becomes "the short regression" by imposing the restriction $\boldsymbol{\beta}_2 = 0$.

• Q: What are the implications of using the wrong model, with omitted variables?

A: We already know the answer, we are imposing a wrong restriction: the restricted estimator, **b***, is biased, but it is more efficient.

Specification Errors: Omitted Variables

- Some easily proved results:

$E[\mathbf{b}_1|\mathbf{X}] = E\,[(\mathbf{X}_1'\mathbf{X}_1)^{-1}\mathbf{X}_1'\,\mathbf{y}|\mathbf{X}] = E\,[(\mathbf{X}_1'\mathbf{X}_1)^{-1}\mathbf{X}_1'\,(\mathbf{X}_1\boldsymbol{\beta}_1 + \mathbf{X}_2\boldsymbol{\beta}_2 + \boldsymbol{\varepsilon})|\mathbf{X}]$

$\quad = \boldsymbol{\beta}_1 + (\mathbf{X}_1'\mathbf{X}_1)^{-1}\mathbf{X}_1'\mathbf{X}_2\boldsymbol{\beta}_2 \neq \boldsymbol{\beta}_1.$

Thus, unless $\mathbf{X}_1'\mathbf{X}_2 = 0$, $\mathbf{b}_1$ is *biased*. The bias can be **huge**. It can reverse the sign of a price coefficient in a "demand equation."

(2) $\text{Var}[\mathbf{b}_1|\mathbf{X}] \leq \text{Var}[\mathbf{b}_{1.2}|\mathbf{X}]$, where $\mathbf{b}_{1.2}$ is the OLS estimator of $\boldsymbol{\beta}_1$ in the long regression (the true model).

Thus, we get a smaller variance when we omit $\mathbf{X}_2$.

Interpretation: Omitting $\mathbf{X}_2$ amounts to using extra information –i.e., $\boldsymbol{\beta}_2 = \mathbf{0}$. We estimate a restricted model! Even if the information is wrong, it reduces the variance.

Specification Errors: Omitted Variables

(3) Mean Squared Error (MSE = RSS/T)

If we use MSE as precision criteria for selecting an estimator, $\mathbf{b}_1$ may be more "precise."

Precision = Mean squared error (MSE)
= Variance + Squared bias.

Smaller variance but positive bias. If bias is small, a practitioner may still favor the short regression.

Note: Suppose $\mathbf{X}_1'\mathbf{X}_2 = \mathbf{0}$. Then the bias goes away. Interpretation, the information is not "right," it is irrelevant: $\mathbf{b}_1$ is the same as $\mathbf{b}_{1.2}$.

Specification Errors: Omitted Variables

Example: We fit an ad-hoc model for U.S. short-term interest rates ($i_{US,t}$) that includes inflation rate ($I_{US,t}$), changes in the USD/EUR (e_t), money growth rate ($m_{US,t}$), and unemployment ($u_{US,t}$), using monthly data from 1975:Jan - 2020:Jul. That is,

$$i_{US,t} = \beta_0 + \beta_1 I_{US,t} + \beta_2 e_t + \beta_3 m_{US,t} + \beta_4 u_{US,t} + \varepsilon_i$$

```
Fger_da <- read.csv("http://www.bauer.uh.edu/rsusmel/4397/FX_USA_GER.csv", head=TRUE,
sep=",")
us_CPI <- Fger_da$US_CPI
us_M1 <- Fger_da$US_M1
us_i <- Fger_da$US_I3M
us_GDP <- Fger_da$US_GDP
ger_CPI <- Fger_da$GER_CPI
us_u <- Fger_da$US_UN
S_ger <- Fger_da$USD_EUR

T <- length(us_CPI)
us_I <- log(us_CPI[-1]/us_CPI[-T])          # US Inflation: (Log) Changes in CPI
us_mg <- log(us_M1[-1]/us_M1[-T])           # US Money Growth: (Log) Changes in M1
e_ger <- log(S_ger[-1]/S_ger[-T])           #  (Log) Changes in USD/EUR
```

Specification Errors: Omitted Variables

Example (continuation):

```
us_i_1 <- us_i[-1]                          # Adjust sample size of untransformed data
us_u_1 <- us_u[-1]                          # Adjust sample size of untransformed data
us_i_0 <- us_i[-T]                          # lagged interest rates, by removing T observation
xx_i <- cbind(us_I ,e_ger, us_mg, us_u_1)   # X matrix
fit_i <- lm(us_i_1 ~ xx_i)
>  summary(fit_i)

Coefficients:
              Estimate Std. Error  t value  Pr(>|t|)
(Intercept)   2.12516    0.52177   4.073    5.34e-05 ***
xx_i_us_I     410.03733  37.17344  11.030   < 2e-16 ***
xx_i_e_ger    8.90564    4.59915   1.936    0.053343 .
xx_i_us_mg    -50.07811  15.04907  -3.328   0.000935 ***        ⇒ significant.
xx_i_us_u_1   0.22673    0.08346   2.717    0.006805 **         ⇒ significant.
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 3.113 on 542 degrees of freedom
Multiple R-squared:  0.2276,    Adjusted R-squared:  0.2219
F-statistic: 39.93 on 4 and 542 DF,  p-value: < 2.2e-16
```

Specification Errors: Omitted Variables

Example (continuation): Now, we include lagged interest rates

```
xx_i <- cbind(us_I ,e_ger, us_mg, us_u_1, us_i_0)        # X matrix with lagged interest rates
fit_i <- lm(us_i_1 ~ xx_i)
> summary(fit_i)
Coefficients:
              Estimate   Std. Error  t value Pr(>|t|)
(Intercept)   0.101007   0.079458   1.271  0.20420
xx_ius_I      16.367138  6.144709   2.664  0.00796 **
xx_ie_ger     3.112901   0.691673   4.501  8.3e-06 ***  ⇒ now, significant.
xx_ius_mg     1.231633   2.284528   0.539  0.59003      ⇒ now, not significant.
xx_ius_u_1    -0.015444  0.012632   -1.223  0.22199     ⇒ now, not significant.
xx_i_us_i_0   0.22673    0.08346    2.717 0.006805 **   ⇒ significant & huge effect on other coeff.
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 3.113 on 542 degrees of freedom
Multiple R-squared:  0.2276,    Adjusted R-squared:  0.2219
```

Note: Lagged i_{US} ($i_{US, t-1}$) is very significant & changes significance of other variables. It may point out to a general misspecification in (**A1**).

Specification Errors: Irrelevant Variables

- Irrelevant variables . Suppose the correct model is

$\mathbf{y} = \mathbf{X}_1\boldsymbol{\beta}_1 + \boldsymbol{\varepsilon}$ –the "short regression," with $\mathbf{X}_1$

But, we estimate

$\mathbf{y} = \mathbf{X}_1\boldsymbol{\beta}_1 + \mathbf{X}_2\boldsymbol{\beta}_2 + \boldsymbol{\varepsilon}$ –the "long regression."

Some easily proved results: Including irrelevant variables just reverse the omitted variables results:

- It increases variance -the cost of not using information.

- But, it does not create biases.

⇒ Since the variables in $\mathbf{X}_2$ are truly irrelevant, then $\boldsymbol{\beta}_2 = \mathbf{0}$,

so $E[\mathbf{b}_{1.2}|\mathbf{X}] = \boldsymbol{\beta}_1$.

Specification Errors: Irrelevant Variables

- A simple example

Suppose the correct model is: $\mathbf{y} = \beta_1 + \beta_2 \mathbf{X}_2 + \boldsymbol{\varepsilon}$

But, we estimate: $\mathbf{y} = \beta_1 + \beta_2 \mathbf{X}_2 + \beta_3 \mathbf{X}_3 + \boldsymbol{\varepsilon}$

- Unbiased: Given that $\beta_3 = 0$ $\Rightarrow E[b_2|X] = \beta_2$

- Efficiency:

$$Var[b_2|X] = \frac{\sigma^2}{\sum(X_{2i} - \bar{X}_2)^2} \times \frac{1}{1 - r^2_{X_2,X_3}} > \frac{\sigma^2}{\sum(X_{2i} - \bar{X}_2)^2}$$

where r_{X_2,X_3} is the correlation coefficient between X_2 and X_3.

Note: These are the results in general. Note that if $\boldsymbol{X}_2$ & $\boldsymbol{X}_3$ are uncorrelated, there will be no loss of efficiency after all.

9