Lecture 6-a
Testing in the CLM & Model Specification

Brooks (4th edition): Chapters 3 & 4

© R. Susmel, 2025 (for private use, not to be posted/shared online). 1

Review: Testing H_0: $R\beta - q = 0$

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

Approach (1): Wald test.

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

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

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

Approach (2): F-test to compare RSS_U & RSS_R.

$$F = \frac{\frac{RSS_R - RSS_U}{(k_U - k_R)}}{\frac{RSS_U}{(T - k_U)}} \sim F_{J,T-k}. \quad \text{(where } J = k_U - k_R\text{)}$$

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

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

Q: Which model should we use?

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**.

© R. Susmel, 2025. Do not post/share online without written permission

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) **Restricted regression**: Run OLS with all the data, with no distinction between categories (Pooled regression). Keep $\boldsymbol{RSS_R}$.

(2) **Unrestricted regression**: Run two separate OLS, one for each category. Keep RSS_1 & RSS_2 ⇒ $\boldsymbol{RSS_U} = RSS_1 + RSS_2$.

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

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

5

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 **27,322** 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
GENDER_F = gender (1 = female)
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

6

© R. Susmel, 2025. Do not post/share online without written permission

Chow Test: Males or Females visit doctors more?

```
Health_Da <-
read.csv("https://www.bauer.uh.edu/rsusmel/4397/german_health.csv",
head=TRUE, sep=",")

x_fem <- Health_Da$Gender_F
 x_age <- Health_Da$age
 x_edu <- Health_Da$educ
 x_hhinc <- Health_Da$hhninc/100
 x_hhkids <- Health_Da$hhkids
 x_married <- Health_Da$married
 x_white_col <- Health_Da$whitecollar
 x_docvis <- Health_Da$docvis

fit_doc_vis <- lm (x_docvis ~  x_age +  x_edu + x_married + x_white_col +
x_hhkids + x_hhinc)
summary(fit_doc_vis)
```

7

Chow Test: Males or Females visit doctors more?

- OLS Estimation for **ALL**. Keep RSS_{ALL} = **858,435** (= **5.606**^2 * **27,315**)

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)	
(Intercept)	2.683700	0.249282	10.766	< 2e-16	***
x_age	0.061810	0.003444	17.947	< 2e-16	***
x_edu	-0.118858	0.015573	**-7.632**	2.38e-14	***
x_married	-0.090716	0.089056	-1.019	0.308	
x_white_col	-0.115412	0.076540	-1.508	0.132	
x_hhkids	-0.492028	0.080014	-6.149	7.89e-10	***
x_hhinc	-0.015429	0.002046	-7.539	4.87e-14	***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Residual standard error: **5.606** on **27,315** degrees of freedom
Multiple R-squared: 0.02949, Adjusted R-squared: 0.02928
F-statistic: 138.3 on 6 and 27315 DF, p-value: < 2.2e-16

Note: We compute RSS_R by pooling, that is, by imposing no gender effect on the coefficients.

$RSS_R = s^2 * (T-k)$ = **5.606**^2 * **27,315** = **858,435**

8

© R. Susmel, 2025. Do not post/share online without written permission

Chow Test: Males or Females visit doctors more?

Run a regression with only **Women** data. Use Allgen to collect relevant data for women only. We will do a for loop and keep data if x_fem is greater than 0.

```
xx <- cbind(x_fem, x_docvis, x_age,  x_edu, x_married, x_white_col, x_hhkids, x_hhinc)
Allgen = NULL         # Initialize empty (to collect variables by one sex (f/m) only)
i <- 1
T <- length(x_fem)
k <- ncol(xx)

for (i in 1:T) {
  if (xx[i,1] > 0) {
    Allgen = rbind(Allgen, xx[i,2:k])
  }
}

y_g <- Allgen[,1]              # Dependent variable: doctor's visits by women only
x_g <- Allgen[,2:(k-1)]
```

9

Chow Test: Males or Females visit doctors more?

```
fit_doc_vis_f <- lm (y_g ~  x_g)
summary(fit_doc_vis_f)
```

• OLS Estimation for **Women** only. Keep RSS_W = **478,894.2**

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)	
(Intercept)	2.999559	0.453506	6.614	3.88e-11	***
x_gx_age	0.049366	0.005719	8.632	< 2e-16	***
x_gx_edu	-0.048141	0.027011	-1.782	0.0747	.
x_gx_married	-0.119853	0.133846	-0.895	0.3706	
x_gx_white_col	-0.006734	0.124768	-0.054	0.9570	
x_gx_hhkids	-0.636619	0.128844	-4.941	7.87e-07	***
x_gx_hhinc	-0.015651	0.003174	-4.932	8.25e-07	***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Residual standard error: **6.052** on 13075 degrees of freedom
Multiple R-squared: 0.01984,   Adjusted R-squared: 0.01939
F-statistic: 44.11 on 6 and 13075 DF, p-value: < 2.2e-16

10

© R. Susmel, 2025. Do not post/share online without written permission

Chow Test: Males or Females visit doctors more?

```
# Use above code, but  change for loop (now, keep data if  x_fem less than 1)

for (i in 1:T) {
  if (xx[i,1] < 1) {
    Allgen = rbind(Allgen, xx[i,2:k])
  }
}

y_g <- Allgen[,1]              # Dependent variable: doctor's visits by women only
x_g <- Allgen[,2:(k-1)]

fit_doc_vis_m <- lm (y_g ~  x_g)
summary(fit_doc_vis_m)
```

11

Chow Test: Males or Females visit doctors more?

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

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)	
(Intercept)	1.801539	0.290792	6.195	5.98e-10	***
x_gx_age	0.067656	0.004421	15.302	< 2e-16	***
x_gx_edu	-0.105462	0.018814	**-5.605**	2.12e-08	***
x_gx_married	0.022278	0.121467	0.183	0.854480	
x_gx_white_col	-0.367075	0.096300	**-3.812**	0.000139	***
x_gx_hhkids	-0.428916	0.102070	-4.202	2.66e-05	***
x_gx_hhinc	-0.015438	0.002629	-5.872	4.40e-09	***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Residual standard error: **5.118** on 14233 degrees of freedom
Multiple R-squared: 0.03602,   Adjusted R-squared: 0.03561

- Chow Test:

$$F = \frac{(RSS_R - [RSS_1 + RSS_2])/k}{(RSS_1 + RSS_2)/(T-2k)} = \frac{(858{,}435 - [372{,}818.1 + 478{,}894.2])/7}{(372{,}818.1 + 478{,}894.2)/(27{,}323 - 14)}$$

= **31.1178**   ⇒ since F(7, 27309) = **2.009925**   ⇒ reject H_0 at 5% level.

12

© R. Susmel, 2025. Do not post/share online without written permission

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.

13

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.

14

© R. Susmel, 2025. Do not post/share online without written permission

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 $\boldsymbol{RSS_R}$.

(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 $\Rightarrow \boldsymbol{RSS_U} = RSS_1 + RSS_2$.

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

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

15

F-Test: Structural Change & Chow Test

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

Q: Did the dot.com bubble (end of 2001, T_{SB} = 348) affect the structure of the FF Model? Sample: January 1973 – December 2024 (T = 624).

Pooled RSS = **2.0826**

Jan 1973 – Dec 2001 RSS = RSS_1 = **1.3481** (T = 347)

Jan 2002 – June 2024 RSS = RSS_2 = **0.7645** (T = 276)

$$F = \frac{[RSS_R - (RSS_1 + RSS_2)]/J}{(RSS_1 + RSS_2)/(T-k)} = \frac{[2.1299 - (1.3481 + 0.7645)]/4}{(1.3481 + 0.7645)/(623 - 2*4)} = \mathbf{1.2598}$$

$\Rightarrow$ Since $F_{4,619,.05}$ = **3.01**, we cannot reject H_0

	Constant	Mkt – rf	SMB	HML	RSS	T
1973-2025	-0.0049*	0.8920*	-0.2141*	-0.0491	**2.1299**	623
1973-2001	-0.0030	0.7987*	-0.21557*	-0.2264	**1.3481**	347
2002 – 2024	-0.0054	0.9414*	-0.2611*	-0.0309&	**0.7645**	276

16

© R. Susmel, 2025. Do not post/share online without written permission

F-Test: Structural Change & Chow Test

Example (continuation): The R package *sctrucchange* estimates the Chow test. (As usual, you need to install package first.)

```
>library(sctrucchange)
> t_sb <- 348
> sctest(ibm_x ~ Mkt_RF + SMB + HML, type = "Chow", point = T_sb)

        Chow test

data:  ibm_x ~ Mkt_RF + SMB + HML
F = 1.2598, p-value = 0.2846
```

Note: For Coca-Cola (KO), the results are quite different

```
> y <- ko_x
> sctest(y ~ Mkt_RF + SMB + HML, type = "Chow", point = T_sb)

        Chow test

data:  y ~ Mkt_RF + SMB + HML
F = 3.6331, p-value = 0.006158
```

17

F-Test: Structural Change & Chow Test

Example: We test if the Oct 1973 (T_{SB} = **108**) oil shock in quarterly GDP growth rates had an structural change on the GDP model.

We model the GDP 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)
```

18

© R. Susmel, 2025. Do not post/share online without written permission

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
```

19

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*).

20

3

© R. Susmel, 2025. Do not post/share online without written permission

Non-nested Models and Tests

• 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. Typical situation, restricted model (Model R) has fewer regressors:

Model U $\boldsymbol{y} = \boldsymbol{X}\beta + \boldsymbol{W}\,\delta + \boldsymbol{\varepsilon}$ (**Unrestriced**)

Model R $\boldsymbol{y} = \boldsymbol{X}\beta + \xi$ (**Restricted**)

Model U becomes Model R under H_0: $\boldsymbol{\delta} = \boldsymbol{0}$. We know how to select a model, using, for example, a Wald test or an F-test.

Example:

Model U: 3-factor FF Model

$$(r_{IBM,t} - r_f) = \beta_0 + \beta_1\,(r_{m,t} - r_f) + \beta_2\,SMB_t + \beta_3\,HML_t + \varepsilon_t$$

Model R: CAPM (imposing H_0: $\beta_2 = \beta_3 = 0$

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

Non-nested Models and Tests

• Sometimes, we have two rival **non-nested models** -i.e., neither is a restricted version of the other. How do we choose a model?

Example:

Model 1 $\boldsymbol{y} = \boldsymbol{X}\beta + \boldsymbol{W}\delta + \boldsymbol{\varepsilon}$

Model 2 $\boldsymbol{y} = \boldsymbol{X}\beta + \boldsymbol{Z}\gamma + \xi$

• There are two common approaches to select a model:

(1) Encompassing test: Form an "unrestricted" model, combining both models. Then, test the two restricted versions of the model against the unrestricted model.

(2) J-test: Test directly one model against the other.

© R. Susmel, 2025. Do not post/share online without written permission

Non-nested Models and Tests: Encompassing

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

We have two models for excess returns, based on the Fama-French factors, M1 & M2:

M1: $(r_{IBM,t} - r_f) = \gamma_0 + \gamma_1 (r_{m,t} - r_f) + \gamma_2 SMB_t + \gamma_3 HML_t + \varepsilon_t$

M2: $(r_{IBM,t} - r_f) = \delta_0 + \delta_1 (r_{m,t} - r_f) + \delta_2 CMA_t + \delta_3 RMW_t + \varepsilon_t$

To select one model, we use an encompassing test. Steps:

(1) Form a *encompassing model* that nests both rival models −M1 & M2.

$(r_{IBM,t} - r_f) = \beta_0 + \beta_1 (r_{m,t} - r_f) + \beta_2 SMB_t + \beta_3 HML_t + \beta_4 CMA_t + \beta_5 RMW_t + \varepsilon_t$ (The **unrestricted Model**, **ME**.)

(2) Do two F-test to the restrictions of each rival model against **ME**.
(i) Test **ME** (Unrestricted Model) against Model 1: H_0: $\beta_4 = \beta_5 = 0$
(ii) Test **ME** (Unrestricted Model) against Model 2 H_0: $\beta_2 = \beta_3 = 0$

Non-nested Models and Tests: Encompassing

• Steps:

(1) Form a *encompassing* model, combining M1 & M2.

$(r_{IBM,t} - r_f) = \beta_0 + \beta_1 (r_{m,t} - r_f) + \beta_2 SMB_t + \beta_3 HML_t + \beta_4 CMA_t + \beta_5 RMW_t + \varepsilon_t$ (The **unrestricted Model**, **ME**.)

(2) Do two F-test to the restrictions of each rival model against **ME**.
(i) Test **ME** (Unrestricted Model) against Model 1: H_0: $\beta_4 = \beta_5 = 0$
(ii) Test **ME** (Unrestricted Model) against Model 2 H_0: $\beta_2 = \beta_3 = 0$

• Suppose we reject the restrictions against one Ml and we cannot reject the restrictions against M2, we are done: We select the M2, that is, the model that the F test do not reject restrictions (M2).

• Q: What happens if the restrictions cannot be rejected? Then, we prefer the model with the lower F statistic for the test of restrictions.

© R. Susmel, 2025. Do not post/share online without written permission

Non-nested Models and Tests: Encompassing

Example: We have:

Model 1 $\quad \boldsymbol{y} = \boldsymbol{X}\beta + \boldsymbol{W}\delta + \boldsymbol{\varepsilon}$

Model 2 $\quad \boldsymbol{y} = \boldsymbol{X}\beta + \boldsymbol{Z}\gamma + \xi$

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

ME: $\quad \boldsymbol{y} = \boldsymbol{X}\beta + \boldsymbol{W}\delta + \boldsymbol{Z}\,\gamma + \boldsymbol{\varepsilon}$

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

F-test for H_0: $\gamma = \mathbf{0}$: 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: $\gamma = \mathbf{0}$ $\Rightarrow$ We have evidence against **Model 1**
If we reject H_0: $\boldsymbol{\delta} = \mathbf{0}$ $\Rightarrow$ 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**: In equilibrium, changes in exchange rates ($\boldsymbol{e}_f$) 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:

$$\boldsymbol{e}_f = \alpha^1 + \beta^1 (\boldsymbol{i}_d - \boldsymbol{i}_f) + \boldsymbol{\varepsilon}^1$$

• **Relative PPP**: 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 DGP consistent with IFE is:

$$\boldsymbol{e}_f = \alpha^2 + \beta^1 (\boldsymbol{I}_d - \boldsymbol{I}_f) + \boldsymbol{\varepsilon}^2$$

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

© R. Susmel, 2025. Do not post/share online without written permission

Non-nested Models and Tests: IFE or PPP?

Example: What drives log changes in exchange rates for the USD/GBP (e_f): $(i_d - i_f)$ or $(I_d - I_f)$?

Model 1 (IFE): $\quad e_f = \alpha^1 + \beta^1 (i_d - i_f) + \varepsilon^1$

Model 2 (PPP): $\quad e_f = \alpha^2 + \beta^2 (I_d - I_f) + \varepsilon^2$

```
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

$$e_f = \alpha + \beta^1 (i_d - i_f) + \beta_2 (I_d - I_f) + \varepsilon \qquad \textbf{(U Model)}$$

```
# 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
```

⇒ cannot reject H_0: $\beta_1 = 0$. (int_dif)

⇒ reject H_0: $\beta_2 = 0$. (inf_dif)

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
```

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

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

- **Testing-based Method 2:** Davidson-MacKinnon (1981)'s ***J*-test**.

We start with two non-nested models. Say,

Model 1: $\boldsymbol{y} = \boldsymbol{X}\beta + \boldsymbol{\varepsilon}$

Model 2: $\boldsymbol{y} = \boldsymbol{Z}\gamma + \xi$

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** ⇒ obtain fitted values: $\boldsymbol{X}\mathbf{b}$.

(2) Add $\boldsymbol{X}\mathbf{b}$ to the list of regressors in Model 2

$$\Rightarrow \boldsymbol{y} = \boldsymbol{Z}\gamma + \lambda\, \boldsymbol{X}\mathbf{b} + \xi$$

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

© R. Susmel, 2025. Do not post/share online without written permission

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

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

(4.1) Estimate **Model 2** ⇒ obtain fitted values: $\boldsymbol{Zc}$.

(4.2) Add $\boldsymbol{Zc}$ to the list of regressors in Model 1:

$$\Rightarrow \boldsymbol{y} = \boldsymbol{X}\beta + \lambda\, \boldsymbol{Zc} + \boldsymbol{\varepsilon}$$

(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.

• "Best situation:" We only reject one H_0: $\lambda = 0$. In this case, it is clear which model to select.

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

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

• 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.

• 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)
("*Best situation*")	$\lambda \neq 0$ and $\lambda = 0$	(**Model 1** is OK)
(3) Both rejected:	$\lambda \neq 0$ and $\lambda \neq 0$	⇒ new model is needed.

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

© R. Susmel, 2025. Do not post/share online without written permission

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

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

Model 1 (IFE): $\boldsymbol{e_f} = \alpha^1 + \beta^1\,(\boldsymbol{i_d} - \boldsymbol{i_f}) + \boldsymbol{\varepsilon}^1$

Model 2 (PPP): $\boldsymbol{e_f} = \alpha^2 + \beta^2\,(\boldsymbol{I_d} - \boldsymbol{I_f}) + \boldsymbol{\varepsilon}^2$

```
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: $\lambda=0$. (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: $\lambda=0$. (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: $\lambda=0$. (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** (an "aggressive test") 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 percentiles 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."

© R. Susmel, 2025. Do not post/share online without written permission

Testing Remarks: A word about α

Practical advise: 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 rejecting 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: We have N = 1,000 tests of H_0. If $\alpha = 0.05$, proportion of false H_0=10%, and the power of the test equals 50% (π = .50), then, **47.4%** of rejections are true H_0 -i.e., "*false positives*."

© R. Susmel, 2025. Do not post/share online without written permission

Testing Remarks: A word about rejecting H_0

Example: We conduct a 1,000 studies of some H_0 (say, H_0: μ = 0)
- Assume the proportion of false H_0 is 10% (**100** false cases).
- Use standard 5% significance level (**45** rejections under H_0).
- Power **50%** (50% correct rejections)

	State of World	
Decision	H_0 true	H_1 true (H_0 false)
Cannot reject H_0	855	50 (Type II error)
Reject H_0	**45** (*Type I error*)	50
	900	**100**

Note: Of the 95 studies which result in a "*statistically significant*" (i.e., p<0.05) result, **45** (**47.4%**) are true H_0 and so are "false positives."

41

Testing Remarks: A word about rejecting H_0

Example (continuation): Now, with same proportion of false H_0 (**10%**) and same α = **5%**, assume the power is **80%** (80% correct rejections of H_0).

	State of World	
Decision	H_0 true	H_1 true (H_0 false)
Cannot reject H_0	855	20 (*Type II error*)
Reject H_0	**45** (*Type I error*)	**80**
	900	**100**

Now, of the 125 studies which result in a "*statistically significant*" (i.e., p<0.05) result, **45** (**36%**) are true H_0 and so are "false positives."

Conclusion: Higher power, lower proportion of false positives.

42

© R. Susmel, 2025. Do not post/share online without written permission

Model Specification: Checking (A1)

OLS Estimation – Assumptions

- CLM Assumptions
  - (**A1**) DGP: $\boldsymbol{y} = \boldsymbol{X}\beta + \boldsymbol{\varepsilon}$ is correctly specified.
  - (**A2**) E[$\boldsymbol{\varepsilon}|\boldsymbol{X}$] = 0
  - (**A3**) Var[$\boldsymbol{\varepsilon}|\boldsymbol{X}$] = $\sigma^2\mathbf{I}_T$
  - (**A4**) $\boldsymbol{X}$ has full column rank –rank($\boldsymbol{X}$)= k–, where $T \geq k$.

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

- We look at (**A1**). We have already studied what happens when we impose restrictions in the DGP: If we impose a true restrictions, estimation is unbiased & more efficient; a false restriction causes bias!

<u>This short lecture</u>: Are we omitting a relevant regressor? Are we including an irrelevant variable? Can we test for omitted variables?

© R. Susmel, 2025. Do not post/share online without written permission

Specification Errors: Omitted & Irrelevant X's

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

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

But, we compute OLS omitting $\boldsymbol{X}_2$, a true driver of $\boldsymbol{y}$. That is,

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

Implication: Restricted estimator **b*** is **biased**, but **more efficient**.

• Irrelevant variables . Suppose the correct model is

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

But, we estimate, ignoring the true restriction $\boldsymbol{\beta}_2 = 0$:

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

Implication: Estimator **b** is **unbiased**, but **inefficient**.

Specification Errors: Omitted & Irrelevant X's

• Omitting relevant variables is considered a very serious problem since it produces a biased (& inconsistent) estimator. In general, inconsistency is a fatal flaw for an estimator.

• For this reason it is very important that (**A1**) does not have omitted variables. A lot of the testing in econometrics is devoted to making sure that there are no omitted variables in a model.

Example: We have used the Wald test to test the CAPM for omitted variables: SMB and HML. In all our tests, we have rejected the restricted model, the CAPM, in favor of the 3-factor F-F model. That is, the CAPM has omitted variables.

© R. Susmel, 2025. Do not post/share online without written permission

Trilogy of Asymptotic Tests: LR, Wald, and LM

• **Wald test**: In practice, to test joint hypothesis H_0: $\boldsymbol{R}\beta = \boldsymbol{q}$, we rely on the asymptotic distribution of the Wald test. We constructed the Wald test based on the unrestricted estimation (OLS), which in the produces **b** and s^2.

Then, the Wald test:

$$W* = (\boldsymbol{Rb} - \boldsymbol{q})' \; \{\boldsymbol{R}[s^2(\boldsymbol{X}'\boldsymbol{X})^{-1}]\boldsymbol{R}\}^{-1} \, (\boldsymbol{Rb} - \boldsymbol{q}) \xrightarrow{d} \chi^2_J.$$

• In more general terms, we construct the Wald test based on the unrestricted estimation, which produces $\hat{\theta}^U$.

• There are two other popular tests that are asymptotically equivalent –i.e., with same asymptotic distribution: the Likelihood Ratio (**LR**) and the Lagrange Multiplier (**LM**) tests.

Trilogy of Asymptotic Tests: LR

• **LR test**: Based on the (log) *Likelihood*. It needs two ML estimations:
- The unrestricted estimation, producing $\hat{\theta}_{ML}$
- The restricted estimation, producing $\hat{\theta}^R$.

Then, the LR test:

$$LR = 2[\log(L(\hat{\theta}_{ML})) - \log(L(\hat{\theta}^R))] \xrightarrow{d} \chi^2_J$$

Note: MLE requires **assuming a distribution**, usually, a normal.

Technical note: The LR test is a **consistent test**. An asymptotic test which rejects H_0 with probability one when the H_1 is true is called a *consistent test*. That is, a consistent test has asymptotic power of 1. The LR test is a consistent test.

LR Test: Are SMB and HML Priced Factors?

Example: We do a LR test to test if the SMB & HML FF factors are significant, using monthly data 1973 – 2025 (T = 624). That is,

$$H_0: \beta_{SMB} = \beta_{HML} = 0$$

We use the function *lrtest* from R package *lmtest*. It assumes normality.

```
library(lmtest)
fit_ibm_ff3 <- lm (ibm_x ~ Mkt_RF + SMB + HML)
fit_ibm_capm <- lm (ibm_x ~ Mkt_RF)
lrtest(fit_ibm_ff3, fit_ibm_capm)

Likelihood ratio test

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

#Df LogLik Df  Chisq Pr(>Chisq)
1   5 884.85
2   3 880.93 -2 7.8329    0.01991 *
```

⇒ p-value is small: Reject H_0

Trilogy of Asymptotic Tests: LM

• **LM Test:** It needs only one estimation: the restricted estimation, that is, imposing H_0: $\boldsymbol{R}\beta - \boldsymbol{q} = \boldsymbol{0}$, producing $\hat{\theta}^R$.

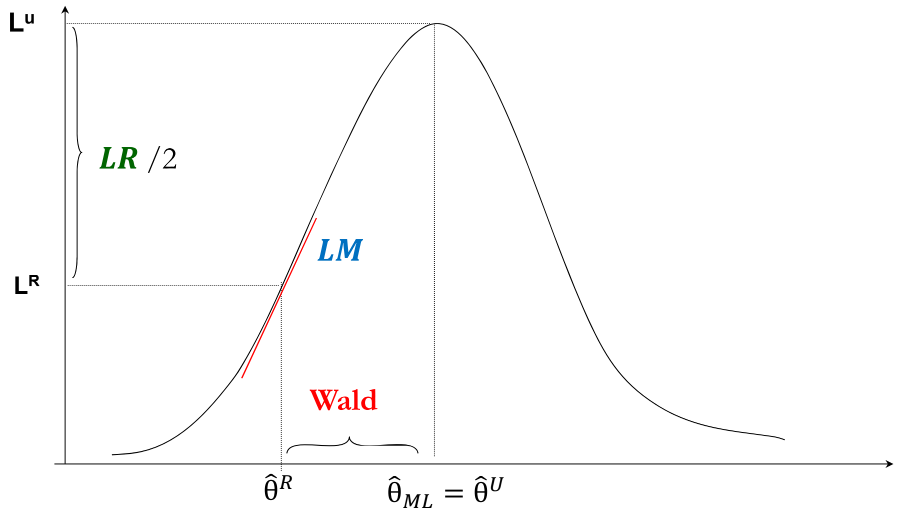
Then, if the restriction is true, then the slope of the objective function (say, the Likelihood) at $\hat{\theta}^R$ should be zero. The slope is called the **Score**, $S(\hat{\theta}^R)$.

• The LM test is based on a Wald test on H_0: $S(\hat{\theta}^R) = 0$.

$$LM = S(\hat{\theta}^R)'[Var(S(\hat{\theta}^R)]^{-1}S(\hat{\theta}^R) \xrightarrow{d} \chi_J^2$$

It turns out that there is a much simpler formulation for the LM test, based on the residuals of the restricted model. We will present this version of the test next.

Trilogy of Asymptotic Tests: LR, Wald, and LM



Remark: Asymptotically equivalent, but, for small T, in general,
$W > LR > LM$.

Testing Model Specification with an LM Test

• We can use the LM test to check for omitted variables, H_0: $\beta = \mathbf{0}$. We have already presented LR & W tests of H_0: $\beta = \mathbf{0}$. Why use an LM test? LM tests only use the restricted estimation, producing $\hat{\theta}^R$.

• The simpler formulation of the LM test is based on the residuals of the restricted model, $\boldsymbol{e}_R$.

Simple intuition. Everything that is omitted from (& belongs to!) a model should appear in the residuals ($\boldsymbol{e}_R$). Suppose the true model is:

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta}_1 + \boldsymbol{Z}\,\boldsymbol{\beta}_2 + \boldsymbol{\varepsilon}$$

But, we omit the J variables, $\boldsymbol{Z}$, implicitly imposing H_0: $\boldsymbol{\beta}_2 = \mathbf{0}$:

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta}_1 + \boldsymbol{\varepsilon} \qquad (\boldsymbol{\varepsilon} = \boldsymbol{Z}\,\boldsymbol{\beta}_2 + \boldsymbol{\varepsilon})$$

LM tests check if $\boldsymbol{e}_R$ can be explained by the J omitted variables $\boldsymbol{Z}$. 52

Testing Model Specification with an LM Test

• We use a simple regression of $\boldsymbol{e}_R$ against $\boldsymbol{Z}$ (dimension JxT) to check for the misspecification.

• LM test steps:

(1) Run restricted model ($\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta}_1 + \boldsymbol{\varepsilon}$). Get restricted residuals, $\boldsymbol{e}_R$.

(2) (Auxiliary Regression). Run the regression of $\boldsymbol{e}_R$ on all the omitted J variables, $\boldsymbol{Z}$, and the k included variables, $\boldsymbol{X}$. In our case:

$$e_{R,i} = \alpha_0 + \alpha_1\, x_{1,i} + \ldots + \alpha_k\, x_{k,i} + \gamma_1\, z_{1,i} + \ldots + \gamma_J\, z_{1,J} + v_i$$

⇒ Keep the R^2 from this regression, R^2_{eR}.

(3) Compute LM-statistic:

$$\text{LM} = T * R^2_{eR} \xrightarrow{d} \chi^2_J.$$

53

Testing Model Specification with an LM Test

Technical Note: We include the original variables in (2), $\boldsymbol{X}$, in the auxiliary regression to get the convenient form for the LM-test.

• The LM Test is very general. It can be used in many settings, for example, to test for nonlinearities, interactions among variables, autocorrelation or heteroscedasticity (discussed later).

Remark: Asymptotically speaking, the LM Test, the LR Test and the Wald Test are equivalent. They have the same limiting distribution, χ^2_J.

With small T, they can have different conclusions. In general, we find:

$$W > LR > LM.$$

That is, the LM test is more conservative (rejects less often) and the Wald test is more aggressive (rejects more often).

54

© R. Susmel, 2025. Do not post/share online without written permission

Testing Model Specification with an LM Test

Example: We use an LM test to check if the standard CAPM for IBM returns omits **SMB** and **HML**. (J = 2)

```
fit_ibm_capm <- lm (ibm_x ~ Mkt_RF)                  # Restricted Model
resid_r <- fit_ibm_capm$residuals                    # extract residuals from R model
fit_lm <- lm (resid_r ~ Mkt_RF + SMB + HML)          # auxiliary regression
> summary(fit_lm)

Coefficients:
             Estimate   Std. Error t value Pr(>|t|)
(Intercept)  0.0003065  0.0023892   0.128  0.89797
Mkt_RF       0.0276474  0.0541081   0.511  0.60956
SMB         -0.2140617  0.0803507  -2.664  0.00792 **
HML         -0.0491383  0.0779484  -0.630  0.52867
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.05866 on 619 degrees of freedom
Multiple R-squared:  0.01249, Adjusted R-squared:  0.007708
F-statistic: 2.611 on 3 and 619 DF,  p-value: 0.05058
```

55

Testing Model Specification with an LM Test

Example (continuation):

```
R2_r <- summary(fit_lm)$r.squared                # extracting  R² from fit_lm
> R2_r
[1] 0.01249

LM_test <- R2_r * T
> LM_test
[1] 7.78127                                      ⇒ LM_test > qchisq (.95,df=2) ⇒ Reject H0.

> qchisq(.95, df = 2)                            # chi-squared (df=2) value at 5% level
[1] 5.991465
p_val <- 1 - pchisq(LM_test, df = 2)             # p-value of LM_test
> p_val
[1] 0.0204331                                    ⇒ p-value is small ⇒ Reject H0.
```

Note: In Lecture 5 we performed the same test with the Wald test (using the F distribution), the p-value was **0.0204**. (This almost exact coincidence is not always the case.)

56

© R. Susmel, 2025. Do not post/share online without written permission