

Econometrics

Practical Session 10

Inference on Linear Combinations of Parameters

Ricardo Gouveia-Mendes

rgouveiamendes@ucp.pt

Spring 2025-26

Católica-Lisbon School of Business and Economics



CATÓLICA
LISBON
BUSINESS & ECONOMICS

Theoretical Wrap-up

Can We Test Restrictions With >1 Parameters?

- We already know how to test **restrictions involving only one parameter**: $H_0 : \beta_j = c$
- But what if we want to test **restrictions involving multiple parameters**? $H_0 : a\beta_j + b\beta_k = c$
- Notice that **defining** $\theta = a\beta_j + b\beta_k$ allows us to rewrite the restriction as $H_0 : \theta = c$

$$t = \frac{\hat{\theta} - c}{s_{\hat{\theta}}} \sim \mathcal{N}(0, 1) \quad \text{CI}_{\theta}^{1-\alpha} = \left\{ \hat{\theta} \pm \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \cdot s_{\hat{\theta}} \right\}$$

- **Problem: $s_{\hat{\theta}}$ is not directly available**

Can We Test Restrictions With >1 Parameters?

- If $\text{cov}(\hat{\beta}_j, \hat{\beta}_k) = 0 \rightarrow s_{\hat{\theta}} = \sqrt{a^2 s_{\hat{\beta}_j}^2 + b^2 s_{\hat{\beta}_k}^2}$
- Alternatively, notice that:

$$\theta = a\beta_j + b\beta_k \iff \beta_j = \frac{\theta}{a} - \frac{b}{a}\beta_k$$

$$Y = \beta_0 + \dots + \left(\frac{\theta}{a} - \frac{b}{a}\beta_k\right)X_j + \beta_k X_k + \dots + u \iff$$

$$Y = \beta_0 + \dots + \theta \frac{X_j}{a} + \beta_k \left(X_k - X_j \cdot \frac{b}{a}\right) + \dots + u$$

- Then, use $\boxtimes$

What About Multiple Linear Restrictions?

- We can test **several linear restrictions at the same time**
- Usually, this means testing: $H_0 : \beta_1 = \beta_2 = \dots = \beta_k = 0$. Example:

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

$$Y = \beta_0 + \beta_1 X_1 + v$$

$$H_0 : \beta_2 = \beta_3 = 0 \quad H_A : \neg H_0$$

What About Multiple Linear Restrictions?

- The statistic is:

$$F = \frac{SSR_* - SSR}{SSR} \frac{N - p}{q} = \frac{R^2 - R_*^2}{1 - R^2} \frac{N - p}{q} \sim F_{q, N-p}$$

- SSR_* and R_*^2 are from the restricted model
- SSR and R^2 are from the unrestricted model
- q is the number of restrictions
- p is the number of parameters in the unrestricted model

Exercises

Stock & Watson – Exercise 7.5

Consider the following regression results, using 2012 data converted into 2007 units using the Consumer Price Index:

$$\widehat{\log(\text{AHE})} = 12.471 + 0.373 \text{ HighSchool} + 0.457 \text{ Male} + 0.011 \text{ Age}$$

(0.049) (0.021) (0.020) (0.001)

$$\text{SER} = 1.023, \quad \bar{R}^2 = 0.0761$$

Consider also a new estimation for the same regression, but using 1993 data also converted to 2007 units:

$$\widehat{\log(\text{AHE})} = 9.32 + 0.301 \text{ HighSchool} + 0.562 \text{ Male} + 0.011 \text{ Age}$$

(0.20) (0.019) (0.047) (0.002)

$$\text{SER} = 1.25, \quad \bar{R}^2 = 0.08$$

Stock & Watson – Exercise 7.5

Was there a statistically significant change in the coefficient on *HighSchool*?

KEY TAKEAWAYS

- We want to test: $H_0 : \theta \equiv \beta_{1,2012} - \beta_{1,1993} = 0$ against $H_A : \theta \neq 0$
- To compute a t -statistic for θ we need to estimate $\text{var}(\hat{\theta}) \equiv \text{var}(\hat{\beta}_{1,2012} - \hat{\beta}_{1,1993})$
- Since the two estimates come from different independent samples:

$$\text{var}(\hat{\theta}) = \text{var}(\hat{\beta}_{1,2012}) + \text{var}(\hat{\beta}_{1,1993}) \Rightarrow s_{\hat{\theta}} = \sqrt{\text{var}(\hat{\beta}_{1,2012}) + \text{var}(\hat{\beta}_{1,1993})} = 0.02832$$

$$t = \frac{\hat{\theta} - 0}{s_{\hat{\theta}}} = 2.542373 > 1.96 = \Phi^{-1}(97.5\%)$$

- We reject H_0 at the 5% significance level: the effect of high school education on average hourly earnings has increased significantly

Stock & Watson – Exercise 7.7

Data were collected from a random sample of 200 home sales from a community in 2013. Let *Price* denote the selling price (in \$1000s), *BDR* denote the number of bedrooms, *Bath* denote the number of bathrooms, *Hsize* denote the size of the house (in square feet), *Lsize* denote the lot size (in square feet), *Age* denote the age of the house (in years), and *Poor* denote a binary variable that is equal to 1 if the condition of the house is reported as “poor.”

Stock & Watson – Exercise 7.7

An estimated regression yields:

$$\widehat{\text{Price}} = 109.7 + 0.567 \text{ BDR} + 26.9 \text{ Bath} + 0.239 \text{ Hsize} + 0.005 \text{ Lsize} \\ + 0.1 \text{ Age} - 56.9 \text{ Poor}, \quad \bar{R}^2 = 0.85, \quad \text{SER} = 45.8$$

(22.1) (1.23) (9.76) (0.021) (0.00072)
(0.23) (12.23)

a) Is the coefficient on *BDR* statistically significantly different from zero?

KEY TAKEAWAYS

$$H_0 : \beta_1 = 0 \quad H_1 : \beta_1 \neq 0$$
$$t = \frac{0.567 - 0}{1.23} = 0.46098 < 1.96 = \Phi^{-1}(97.5\%)$$

We do not reject H_0 , so the coefficient on *BDR* is not statistically significantly different from zero.

Stock & Watson – Exercise 7.7

b) Typically, four-bedroom houses sell for more than three-bedroom houses. Is this consistent with your answer to a) and with the regression more generally?

KEY TAKEAWAYS

- The coefficient on *BDR* is positive, pointing to a higher price the higher the number of bedrooms
- However it is not statistically significant, so we cannot conclude that four-bedroom houses sell for more than three-bedroom houses

c) A homeowner purchases 2500 square feet from an adjacent lot. Construct a 95% confident interval for the change in the value of her house.

KEY TAKEAWAYS

- The estimator can be written as: $\theta = 2500 \beta_4$
- To build a confidence interval we need: $CI_{\theta}^{95\%} = \{\hat{\theta} \pm s_{\hat{\theta}} \cdot \Phi^{-1}(97.5\%)\}$

$$\hat{\theta} = 2500 \cdot 0.005 = 12.5 \quad s_{\hat{\theta}} = \sqrt{2500^2 \cdot s_{\hat{\beta}_4}^2} = 2500 \cdot s_{\hat{\beta}_4} = 1.8$$

$$CI_{\theta}^{95\%} = [8.972, 16.028]$$

Stock & Watson – Exercise 7.7

d) Lot size is measured in square feet. Do you think that another scale might be more appropriate? Why or why not?

KEY TAKEAWAYS

- The scale of the variables should make easy to read the regression results
- If the lot size were measured in thousands of square feet, the estimate coefficient would be 5 instead of 0.005

Stock & Watson – Exercise 7.7

e) The F -statistic for omitting BDR and Age from the regression is $F = 2.38$. Are the coefficients on BDR and Age statistically different from zero at the 10% level?

KEY TAKEAWAYS

- $F_{2,193}^{-1}(90\%) = 2.33 < 2.38$
- So we reject the null hypothesis that the coefficients on BDR and Age are zero at the 10% level
- The coefficients are jointly significant