# Introducció a l'Econometria
## Capítol 5

**Ezequiel Uriel Jiménez**
Universitat de València

València, 2019

# 5 Anàlisi de regressió múltiple amb informació qualitativa

# 5.1 Introducció d'informació qualitativa en els models economètrics



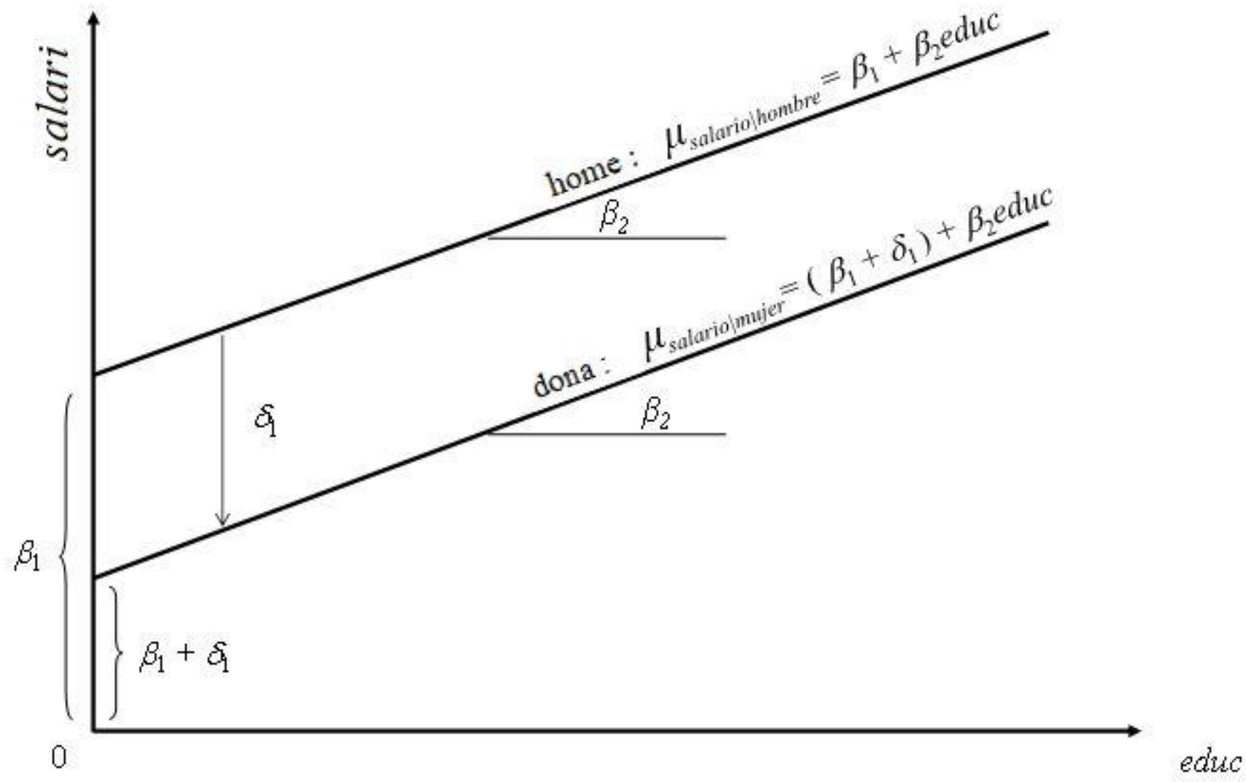


**FIGURA 5.1. Mateix pendent, terme independent diferent.**

# 5.2 Una sola variable fictícia independent



**EXEMPLE 5.1 Existeix discriminació salarial per a la dona a Espanya? (fitxer *wage02sp*)**

$$\ln(wage) = \beta_1 + \delta_1\, female + \beta_2 educ + u$$

$$\ln(wage) = \underset{(0.026)}{1.731} - \underset{(0.022)}{0.307}\ female + \underset{(0.0025)}{0.0548}\, educ$$

$$SQR = 393 \quad R^2 = 0.243 \quad n = 2000$$

$$H_0 : \delta_1 = 0$$
$$H_1 : \delta_1 < 0$$

$$t = \frac{-0.3070}{0.0216} = -14.26$$

La diferència percentual en el salari per hora entre homes i dones és

$$= \ 100 \times (e^{0.307} - 1) = 35.9\%$$

# 5.2 Una sola variable fictícia independent



**EXEMPLE 5.2 Anàlisi de la relació entre la capitalització de mercat i el valor comptable: el paper de l'IBEX-35 (fitxer *bolmad11*)**

$$\ln(marketcap) = \beta_1 + \delta_1 ibex35 + \beta_2 \ln(bookvalue) + u$$

$$\ln(marketcap) = \underset{(0.243)}{1.784} + \underset{(0.179)}{0.690}\, ibex35 + \underset{(0.037)}{0.675} \ln(bookvalue)$$

$$SQR = 35.672 \quad R^2 = 0.893 \quad n = 92$$

$$H_0 : \delta_1 = 0$$
$$H_1 : \delta_1 > 0$$

$$t = \frac{0.690}{0.179} = 3.85$$

Diferència percentual= $100 \times (e^{0.690} - 1) = 99.4\%$

$$H_0 : \beta_2 = 0$$
$$H_1 : \beta_2 \neq 0$$

$$t = \frac{0.675}{0.037} = 18$$

# 5.2 Una sola variable fictícia independent

**EXEMPLE 5.3 Gasten més en peix les persones que viuen en zones urbanes que les que viuen en zones rurals? (fitxer *demand*)**

$$\ln(fish) = \beta_1 + \delta_1 urban + \beta_2 \ln(inc) + u$$

$$\ln(fish) = \underset{(0.511)}{-6.375} + \underset{(0.055)}{0.140} urban + \underset{(0.070)}{1.313} \ln(inc)$$

$$SCR = 1.131 \quad R^2 = 0.904 \quad n = 40$$

$$H_0 : \delta_1 = 0$$
$$H_1 : \delta_1 > 0$$

$$t = \frac{0.140}{0.055} = 2.55$$

# 5.3 Categories múltiples per a un atribut



***El parany de les variables fictícies***

***Exemple***

$$\ln(wage) = \beta_1 + \theta_0 small + \theta_1 medium + \theta_2 large + \beta_2 educ + u$$

$$\mathbf{X} = \begin{bmatrix} 1 & 1 & 0 & 0 & educ_1 \\ 1 & 1 & 0 & 0 & educ_2 \\ 1 & 0 & 1 & 0 & educ_3 \\ 1 & 0 & 1 & 0 & educ_4 \\ 1 & 0 & 0 & 1 & educ_5 \\ 1 & 0 & 0 & 1 & educ_6 \end{bmatrix}$$

**Solució:**

$$\ln(wage) = \beta_1 + \theta_1 medium + \theta_2 large + \beta_2 educ + u$$

$$\ln(wage) = \theta_0 small + \theta_1 medium + \theta_2 large + \beta_2 educ + u$$

# 5.3 Categories múltiples per a un atribut

**EXEMPLE 5.4 Influeix la mida de l'empresa en la determinació dels salaris? (fitxer *wage02sp*)**

$$\ln(wage) = \beta_1 + \theta_1 medium + \theta_2 large + \beta_2 educ + u$$

$$\ln(wage) = \underset{(0.027)}{1.566} + \underset{(0.025)}{0.281} medium + \underset{(0.024)}{0.162} large + \underset{(0.0025)}{0.0480} educ$$

$$SQR = 406 \quad R^2 = 0.218 \quad n = 2000$$

$$H_0 : \theta_1 = \theta_2 = 0$$

$$H_1 : H_0 \text{ no és certa}$$

$$\ln(wage) = \beta_1 + \beta_2 educ + u$$

$$\ln(wage) = \underset{(0.026)}{1.657} + \underset{(0.0026)}{0.0525} educ$$

$$SQR = 433 \quad R^2 = 0.166 \quad n = 2000$$

$$F = \frac{\left[SQR_R - SQR_{NR}\right]/q}{SQR_{NR}/(n-k)} = \frac{\left[433 - 406\right]/2}{406/(2000-4)} = 66.4$$

# 5.3 Categories múltiples per a un atribut



**EXEMPLE 5.5 En el cas de Lydia E. Pinkham, són significatives les variables temporals fictícies de forma individual i conjunta? (fitxer *pinkham*)**

$$sales_t = \beta_1 + \beta_2 advexp_t + \beta_3 sales_{t-1} + \beta_4 d1_t + \beta_5 d2_t + \beta_6 d3_t + u_t$$

$$sales_t = \underset{(96.3)}{254.6} + \underset{(0.136)}{0.5345}\, advexp_t + \underset{(0.0814)}{0.6073}\, sales_{t-1} - \underset{(89)}{133.35}\, d1_t + \underset{(67)}{216.84}\, d2_t - \underset{(67)}{202.50}\, d3t$$

$$R^2 = 0.929 \quad n = 53$$

$$\begin{cases} H_0 : \theta_i = 0 \\ H_1 : \theta_i \neq 0 \end{cases} \qquad i = 1, 2, 3$$

$$t_{\hat{\theta}_1} = \frac{-133.35}{89} = -1.50 \qquad t_{\hat{\theta}_2} = \frac{216.84}{67} = 3.22 \qquad t_{\hat{\theta}_3} = \frac{-202.50}{67} = -3.02$$

$$\begin{cases} H_0 : \theta_1 = \theta_2 = \theta_3 = 0 \\ H_1 : H_0 \text{ no és certa} \end{cases}$$

$$F = \frac{(R_{NR}^2 - R_R^2)/q}{(1 - R_{NR}^2)/(n-k)} = \frac{(0.9290 - 0.8770)/3}{(1 - 0.9290)/(53-6)} = 11.47$$

**EXEMPLE 5.6 La influència de gènere i durada de la jornada de treball en la determinació dels salaris (fitxer *wage06sp*)**

$$\ln(wage) = \beta_1 + \delta_1 female + \phi_1 partime + \beta_2 educ + u$$

$$\ln(wage) = \underset{(0.026)}{2.006} - \underset{(0.021)}{0.233}\, female - \underset{(0.027)}{0.087}\, partime + \underset{(0.0023)}{0.0531} educ$$

$$SQR = 365 \quad R^2 = 0.235 \quad n = 2000$$

**EXEMPLE 5.7 Anàlisi de l'absentisme laboral a l'empresa Buenosaires (fitxer *absent*)**

$$absent = \beta_1 + \delta_1 bluecoll + \phi_1 male + \beta_2 age + \beta_3 tenure + \beta_4 wage + u$$

$$absent = \underset{(1.640)}{12.444} + \underset{(0.669)}{0.968} bluecoll + \underset{(0.712)}{2.049} male - \underset{(0.047)}{0.037} age - \underset{(0.065)}{0.151} tenure - \underset{(0.007)}{0.044} wage$$

$$SQR = 161.95 \quad R^2 = 0.760 \quad n = 48$$

$$H_0: \delta_1 = 0 \qquad H_1: \delta_1 \neq 0$$

$$H_0: \delta_1 = 0 \qquad H_1: \delta_1 > 0$$

$$t = \frac{0.968}{0.669} = 1.45$$

$$H_0: \varphi_1 = 0 \qquad H_1: \varphi_1 \neq 0$$

$$t = \frac{2.049}{0.712} = 2.88$$

**EXEMPLE 5.8 Mida de l'empresa i gènere en la determinació del salari (fitxer *wage02sp*)**

$$\ln(wage) = \beta_1 + \delta_1 female + \theta_1 medium + \theta_2 large + \beta_2 educ + u$$

$$H_0 : \delta_1 = \theta_1 = \theta_2 = 0$$
$$H_1 : H_0 \text{ no és certa}$$

$$\ln(wage) = \underset{(0.026)}{1.639} - \underset{(0.021)}{0.327}\, female + \underset{(0.023)}{0.308}\, medium + \underset{(0.023)}{0.168}\, large + \underset{(0.0024)}{0.0499}\, educ$$
$$SQR = 361 \quad R^2 = 0.305 \quad n = 2000$$

$$F = \frac{\left[SQR_R - SQR_{NR}\right]/q}{SQR_{NR}/(n-k)} = \frac{\left[433-361\right]/3}{361/(2000-5)} = 133$$

# 5.5 Les interaccions que impliquen variables fictícies



**EXEMPLE 5.9 És la interacció entre les dones i el treball a temps parcial significativa? (fitxer *wage06sp*)**

$$\ln(wage) = \beta_1 + \delta_1 female + \phi_1 partime + \varphi_1 female \times partime + \beta_2 educ + u$$

$$\ln(wage) = \underset{(0.026)}{2.007} - \underset{(0.022)}{0.259}\ female - \underset{(0.047)}{0.198}\ partime + \underset{(0.058)}{0.167}\ female \times partime + \underset{(0.0024)}{0.0538} educ$$

$$SQR = 363 \quad R^2 = 0.238 \quad n = 2000$$

$$H_0 : \varphi_1 = 0 \qquad H_1 : \varphi_1 \neq 0$$

$$t = \frac{0.167}{0.058} = 2.89$$

# 5.5 Les interaccions que impliquenn variables fictícies



**EXEMPLE 5.10 Discriminen les empreses xicotetes a les dones més, o menys, que les empreses grans? (fitxer *wage02sp*)**

$$\ln(wage) = \beta_1 + \delta_1 \, female + \theta_1 medium + \theta_2 large$$
$$+ \varphi_1 \, female \times medium + \varphi_2 \, female \times large + \beta_2 educ + u$$

$$\ln(wage) = \underset{(0.027)}{1.624} - \underset{(0.034)}{0.262} \; female + \underset{(0.028)}{0.361} medium + \underset{(0.027)}{0.179} large$$
$$- \underset{(0.050)}{0.159} \; female \times medium - \underset{(0.051)}{0.043} \; female \times large + \underset{(0.0024)}{0.0497} educ$$
$$SQR = 359 \quad R^2 = 0.308 \quad n = 2000$$

$$H_0 : \varphi_1 = \varphi_2 = 0$$
$$H_1 : H_0 \text{ no és certa}$$

$$F = \frac{\left[SQR_R - SQR_{NR}\right]/q}{SQR_{NR}/(n-k)} = \frac{\left[361 - 359\right]/2}{359/(2000-7)} = 5.55$$

# 5.5 Les interaccions que impliquen variables fictícies



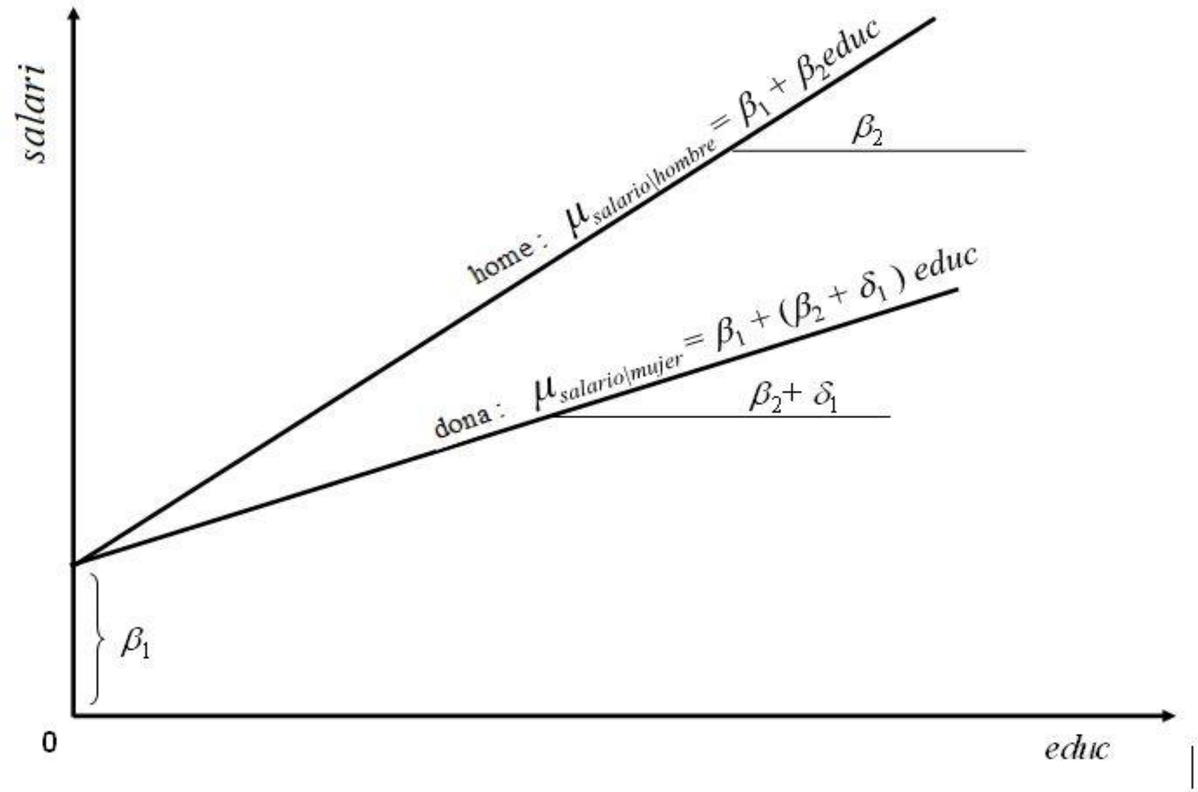


**FIGURA 5.2. Diferent pendent, mateix terme independent.**

# 5.5 Les interaccions que impliquen variables fictícies



**EXEMPLE 5.11 És el rendiment de l'educació per als homes més que per a les dones? (fitxer *wage02sp*)**

$$wage = \beta_1 + \beta_2 educ + \delta_1 female \times educ + u$$

$$\ln(wage) = \underset{(0.025)}{1.640} + \underset{(0.0026)}{0.0632}\, educ - \underset{(0.0021)}{0.0274}\, educ \times female$$

$$SQR = 400 \quad R^2 = 0.229 \quad n = 2000$$

$$H_0 : \delta_1 = 0$$
$$H_1 : \delta_1 < 0$$

$$t = -\frac{0.0274}{0.0021} = -12.81$$

# 5.6 Contrast de canvi estructural



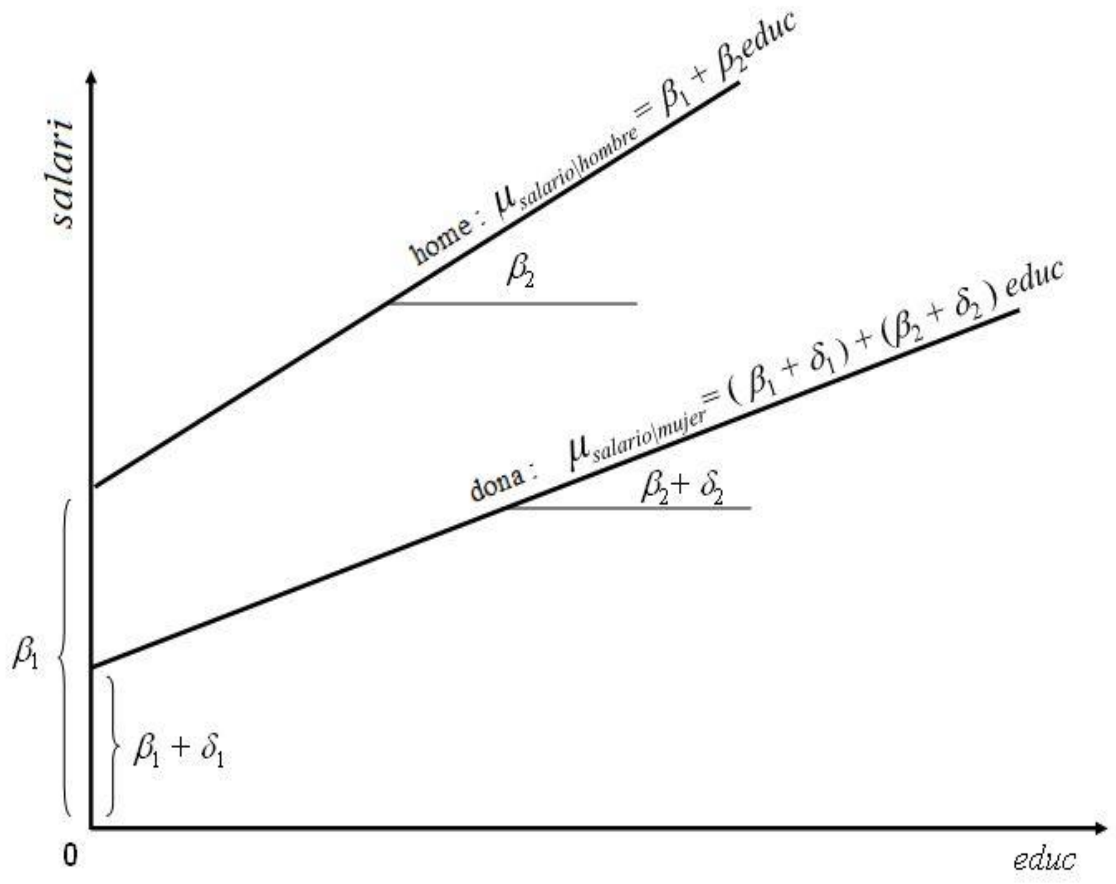


**FIGURA 5. 3. Pendent diferent, diferent terme independent.**

# 5.6 Contrast de canvi estructural

**EXEMPLE 5.12 És l'equació de salaris vàlida tant per a homes com per a dones? (fitxer *wage02sp*)**

$$wage = \beta_1 + \delta_1\, female + \beta_2 educ + \delta_2\, female \times educ + u$$

$$H_0 : \delta_1 = \delta_2 = 0$$
$$H_1 : H_0 \text{ no és certa}$$

$$\ln(wage) = \underset{(0.030)}{1.739} - \underset{(0.0546)}{0.3319}\, female + \underset{(0.0030)}{0.0539}\, educ - \underset{(0.0054)}{0.0027}\, educ \times female$$

$$SQR = 393 \quad R^2 = 0.243 \quad n = 2000$$

$$\ln(wage) = \underset{(0.026)}{1.657} + \underset{(0.0026)}{0.0525}\, educ$$

$$SQR = 433 \quad R^2 = 0.166 \quad n = 2000$$

$$F = \frac{\left[SQR_R - SQR_{NR}\right]/q}{SQR_{NR}/(n-k)} = \frac{\left[433-393\right]/2}{393/(2000-4)} = 102$$

# 5.6 Contrast de canvi estructural

**EXEMPLE 5.13 Tenen els consumidors urbans el mateix patró de comportament que els rurals pel que fa a la despesa en peix? (fitxer *demand*)**

$$\ln(fish) = \beta_1 + \delta_1 urban + \beta_2 \ln(inc) + \delta_2 \ln(inc) \times urban + u$$

$$H_0 : \delta_1 = \delta_2 = 0$$
$$H_1 : H_0 \text{ no és certa}$$

$$\ln(fish) = \beta_1 + \beta_2 \ln(inc) + u$$

$$\ln(fish) = \underset{(0.627)}{-6.551} + \underset{(1.095)}{0.678}\, urban + \underset{(0.087)}{1.337} \ln(inc) - \underset{(0.152)}{0.075} \ln(inc) \times urban$$
$$SQR = 1.123 \quad R^2 = 0.904 \quad n = 40$$

$$\ln(fish) = \underset{(0.542)}{-6.224} + \underset{(0.075)}{1.302} \ln(inc)$$
$$SQR = 1.325 \quad R^2 = 0.887 \quad n = 40$$

$$F = \frac{\left[SQR_R - SQR_{NR}\right]/q}{SQR_{NR}/(n-k)} = \frac{\left[1.325 - 1.123\right]/2}{1.123/(40-4)} = 3.24$$

# 5.6 Contrast de canvi estructural



**EXEMPLE 5.14 Ha canviat l'estructura productiva de les regions espanyoles? (fitxer *prodsp* )**

$$\ln(q) = \gamma_1 + \alpha_1 \ln(k) + \beta_1 \ln(l) + \gamma_2 y2008 + \alpha_2 y2008 \times \ln(k) + \beta_2 y2008 \times \ln(l) + u$$

$$\varepsilon_{Q/K(1995)} = \frac{\partial \ln(Q)}{\partial \ln(K)} = \alpha_1 \qquad \varepsilon_{Q/K(2008)} = \frac{\partial \ln(Q)}{\partial \ln(K)} = \alpha_1 + \alpha_2$$

$$\varepsilon_{Q/K(1995)} = \frac{\partial \ln(L)}{\partial \ln(K)} = \beta_1 \qquad \varepsilon_{Q/K(2008)} = \frac{\partial \ln(L)}{\partial \ln(K)} = \beta_1 + \beta_2$$

$$PEF(1995) = \gamma_1 \qquad PEF(2008) = \gamma_1 + \gamma_2$$

$$H_0 : \gamma_2 = \alpha_2 = \beta_2 \qquad H_1 : H_0 \text{ no és certa}$$

$$\ln(q) = \gamma_1 + \alpha_1 \ln(k) + \beta_1 \ln(l) + u$$

$$\text{Model no restringit : } \ln(gva) = \underset{(0.916)}{0.0559} + \underset{(0.185)}{0.6743} \ln(captot) + \underset{(0.185)}{0.3291} \ln(labour)$$

$$- \underset{(2.32)}{0.1088}\, y2008 + \underset{(0.419)}{0.0154}\, y2008 \times \ln(captot) - \underset{(0.418)}{0.0094}\, y2008 \times \ln(labour)$$

$$R^2 = 0.99394 \quad n = 34$$

$$\text{Model restringit } \ln(gva) = -\underset{(0.200)}{0.0690} + \underset{(0.036)}{0.6959} \ln(captot) + \underset{(0.042)}{0.311} \ln(labour) \qquad R^2 = 0.99392 \quad n = 34$$

$$F = \frac{(R_{NR}^2 - R_R^2)/q}{(1 - R_{NR}^2)/(n-k)} = \frac{(0.99394 - 0.99392)/3}{(1 - 0.99394)/(34-6)} = 0.0308$$

# 5.6 Contrast de canvi estructural

**EXEMPLE 5.15 Una altra forma d'abordar la qüestió de la determinació dels salaris per criteri de gènere (fitxer *wage02sp*)**

Ecuació per a la dona

$$\ln(wage) = \beta_{11} + \beta_{21} educ + u$$

$$\ln(wage) = \underset{(0.042)}{1.407} + \underset{(0.0041)}{0.0566}\, educ$$

$$SQR = 104 \quad R^2 = 0.236 \quad n = 617$$

Ecuació per a l'home

$$\ln(wage) = \beta_{12} + \beta_{22} educ + u$$

$$\ln(wage) = \underset{(0.031)}{1.739} + \underset{(0.0032)}{0.0539}\, educ$$

$$SQR = 289 \quad R^2 = 0.175 \quad n = 1383$$

$$F = \frac{\left[SQR_P - (SQR_F + SQR_M)\right]/k}{SQR_F + SQR_M)/(n-2k)} = \frac{\left[433-(104+289)\right]/2}{(104+289)/(2000-2\times 2)} = 102$$

L'estadístic *F* ha de ser, i ho és, igual al de l'Exemple 5.12.

# 5.6 Contrast de canvi estructural



**EXEMPLE 5.16 El model de determinació dels salaris és el mateix per a diferents mides d'empresa? (fitxer *wage02sp*)**

$$\begin{aligned} xicoteta &: \ln(wage) = \beta_{11} + \delta_{11}\, female + \beta_{21} edu + u \\ mitjana &: \ln(wage) = \beta_{12} + \delta_{12}\, female + \beta_{22} edu + u \\ gran &: \ln(wage) = \beta_{13} + \delta_{13}\, female + \beta_{23} edu + u \end{aligned}$$

$$H_0 : \begin{cases} \beta_{11} = \beta_{12} = \beta_{13} \\ \delta_{11} = \delta_{12} = \delta_{13} \\ \beta_{21} = \beta_{22} = \beta_{23} \end{cases} \qquad H_1 : \text{No } H_0$$

$$xicoteta \quad \widehat{\ln(wage)} = \underset{(0.034)}{1.706} - \underset{(0.031)}{0.249}\, female + \underset{(0.0038)}{0.0396}\, educ \qquad SQR = 121 \quad R^2 = 0.160 \quad n = 801$$

$$mitjana \quad \widehat{\ln(wage)} = \underset{(0.051)}{1.934} - \underset{(0.039)}{0.422}\, female + \underset{(0.0046)}{0.0548}\, educ \qquad SQR = 123 \quad R^2 = 0.302 \quad n = 590$$

$$gran \quad \widehat{\ln(wage)} = \underset{(0.046)}{1.749} - \underset{(0.039)}{0.303}\, female + \underset{(0.0044)}{0.0554}\, educ \qquad SQR = 114 \quad R^2 = 0.273 \quad n = 609$$

$$F = \frac{\left[SQR_P - (SQR_S + SQR_M + SQR_L)\right]/2k}{(SQR_S + SQR_M + SQR_L)/(n-3k)} = \frac{\left[393 - (121 + 123 + 114)\right]/6}{(121 + 123 + 114)/(2000 - 3 \times 3)} = 32.5$$

# 5.6 Contrast de canvi estructural

**EXEMPLE 5.17 És el model Pinkham vàlid per als quatre períodes? (fitxer *pinkham*)**

1907-1914 $sales_t = \beta_{11} + \beta_{21} advexp_t + \beta_{31}\, sales_{t-1} + u_t$

1915-1925 $sales_t = \beta_{12} + \beta_{22} advexp_t + \beta_{32}\, sales_{t-1} + u_t$

1926-1940 $sales_t = \beta_{13} + \beta_{23} advexp_t + \beta_{33}\, sales_{t-1} + u_t$

1941-1960 $sales_t = \beta_{14} + \beta_{24} advexp_t + \beta_{34}\, sales_{t-1} + u_t$

$$H_0 : \begin{cases} \beta_{11} = \beta_{12} = \beta_{13} = \beta_{14} \\ \beta_{21} = \beta_{22} = \beta_{23} = \beta_{24} \\ \beta_{31} = \beta_{32} = \beta_{33} = \beta_{34} \end{cases} \qquad H_1 : \text{No } H_0$$

$$sales_t = \beta_1 + \beta_2 advexp_t + \beta_3 sales_{t-1} + u_t$$

1907-1914 $\widehat{sales}_t = \underset{(603)}{64.84} + \underset{(1.025)}{0.9149}\, advexp + \underset{(0.425)}{0.4630}\, sales_{t-1}$ $SCR = 36017$ $n = 7$

1915-1925 $\widehat{sales}_t = \underset{(190)}{221.5} + \underset{(0.557)}{0.1279}\, advexp + \underset{(0.300)}{0.9319}\, sales_{t-1}$ $SCR = 400605$ $n = 11$

1926-1940 $\widehat{sales}_t = \underset{(112)}{446.8} + \underset{(0.115)}{0.4638}\, advexp + \underset{(0.0827)}{0.4445}\, sales_{t-1}$ $SCR = 201614$ $n = 15$

1941-1960 $\widehat{sales}_t = \underset{(134)}{-182.4} + \underset{(0.241)}{1.6753}\, advexp + \underset{(0.111)}{0.3042}\, sales_{t-1}$ $SCR = 187332$ $n = 20$

$$\widehat{sales}_t = \underset{(95.7)}{138.7} + \underset{(0.156)}{0.3288}\, advexp + \underset{(0.0915)}{0.7593}\, sales_{t-1} \qquad SCR = 2527215 \quad n = 53$$

$$F = \frac{\left[SQR_P - (SQR_1 + SQR_2 + SQR_3 + SQR_4)\right]/3k}{(SQR_1 + SQR_2 + SQR_3 + SQR_4)/(n-4k)}$$

$$= \frac{\left[2527215 - (36017 + 400605 + 201614 + 187332)\right]/9}{(36017 + 400605 + 201614 + 187332)/(53 - 4\times 3)} = 9.16$$