SOLUCION Y GUIA PARA LA CORRECCION DEL EXAMEN DE ECONOMETRIA I.

December 18, 2001

1 SOLUCIÓN EXAMEN ECONOMETRÍA I, SEPTIEMBRE 1999

1.1 PROBLEMA 1

a)

$$\begin{bmatrix} \widehat{\alpha} \\ \widehat{\delta} \end{bmatrix} = (\mathbf{x}'\mathbf{x})^{-1}\mathbf{x}'\mathbf{y}$$

$$\mathbf{x}'\mathbf{x} = \begin{bmatrix} n & \sum z_i \\ \sum z_i & \sum z_i^2 \end{bmatrix} = \begin{bmatrix} 800 & 330 \\ 330 & 144 \end{bmatrix}$$

$$\mathbf{x}'\mathbf{y} = \begin{bmatrix} \sum y_i \\ \sum z_i y_i \end{bmatrix} = \begin{bmatrix} 2672 \\ 1108 \end{bmatrix}$$

$$\begin{bmatrix} \widehat{\alpha} \\ \widehat{\delta} \end{bmatrix} = \frac{1}{800 * 144 - 330^2}\begin{bmatrix} 144 & -330 \\ -330 & 800 \end{bmatrix}\begin{bmatrix} 2672 \\ 1108 \end{bmatrix} = \frac{1}{6300}\begin{bmatrix} 19128 \\ 4640 \end{bmatrix} = \begin{bmatrix} 3.0362 \\ 0.7365 \end{bmatrix}$$

b) Nótese que las perturbaciones son heterocedásticas, y por tanto

$$var\left(\begin{bmatrix} \widehat{\alpha} \\ \widehat{\delta} \end{bmatrix}\right) = (\mathbf{x}'\mathbf{x})^{-1}\sum_i \widehat{u}_i^2 x_i x_i' (\mathbf{x}'\mathbf{x})^{-1} \neq (\mathbf{x}'\mathbf{x})^{-1}$$

donde

$$\sum_i \widehat{u}_i^2 x_i x_i' = \begin{bmatrix} \sum_i \widehat{u}_i^2 & \sum_i \widehat{u}_i^2 z_i \\ \sum_i \widehat{u}_i^2 z_i & \sum_i \widehat{u}_i^2 z_i^2 \end{bmatrix}$$

$$\begin{aligned} var\left(\begin{bmatrix} \widehat{\alpha} \\ \widehat{\delta} \end{bmatrix}\right) &= \frac{1}{6300}\begin{bmatrix} 144 & -330 \\ -330 & 800 \end{bmatrix}\begin{bmatrix} 660 & 309 \\ 309 & 160800 \end{bmatrix}\frac{1}{6300}\begin{bmatrix} 144 & -330 \\ -330 & 800 \end{bmatrix} = \\ &= \frac{1}{6300^2}\begin{bmatrix} -6930 & -8304 \\ 29400 & 26030 \end{bmatrix}\begin{bmatrix} 144 & -330 \\ -330 & 800 \end{bmatrix} = \\ &= \frac{1}{6300^2}\begin{bmatrix} 1742400 & -4356300 \\ -4356300 & 1112200 \end{bmatrix} = \\ &= \begin{bmatrix} 0.0439 & -0.1098 \\ -0.1098 & 0.28022 \end{bmatrix} \end{aligned}$$

Nota:
En caso de que $\widehat{\delta}$ se estimase como:

$$\begin{aligned} \widetilde{y}_i &= \delta \widetilde{z}_i + u_i \\ \widehat{\delta} &= \frac{\sum \widetilde{z}_i \widetilde{y}_i}{\sum \widetilde{z}_i^2} \end{aligned}$$

donde $\widetilde{y}_i$ y $\widetilde{z}_i$ son las desviaciones respecto de la media, $\widetilde{y}_i = y_i - \overline{y}$, $\widetilde{z}_i = z_i - \overline{z}$.
la $\widehat{var}\left(\widehat{\delta}\right)$ se obtiene como:

$$
\begin{aligned}
\widehat{var}\left(\widehat{\delta}\right) &= \left(\sum \widetilde{z}_i^2\right)^{-1}\left(\sum \widehat{u}_i \widetilde{z}_i^2\right)\left(\sum \widetilde{z}_i^2\right)^{-1} = \\
&= \frac{\sum \widehat{u}_i^2 \widetilde{z}_i^2}{\left(\sum \widetilde{z}_i^2\right)^2} = \frac{\sum \widehat{u}_i^2 \left(z_i - \overline{z}\right)^2}{\left[\sum \left(z_i - \overline{z}\right)^2\right]^2} = \\
&= \frac{\sum \widehat{u}_i^2 \left(z_i^2 - 2z_i\overline{z} + \overline{z}^2\right)}{\left(\sum z_i^2 - n\overline{z}^2\right)^2} = \\
&= \frac{\sum \widehat{u}_i^2 z_i^2 - 2\overline{z}\sum \widehat{u}_i^2 z_i + \overline{z}^2 \sum \widehat{u}_i^2}{\left(\sum z_i^2 - n\overline{z}^2\right)^2} = \\
&= \frac{160 - 2\frac{330}{800}309 + \left(\frac{330}{800}\right)^2 660}{\left[144 - 800\left(\frac{330}{800}\right)^2\right]^2} = \\
&= \frac{160 - 254.925 + 272.25}{(144136.125)^2} = 0.28022
\end{aligned}
$$

$$
t(\widehat{\delta}) = \frac{0.7365}{\sqrt{0.28022}} = 1.39
$$

$$
t(\widehat{\delta}) \sim N(0,1) \quad bajo \quad H_0 : \delta = 0
$$

$$
p - valor = 16.5\% \Longrightarrow no\ rechazamos\ H_0
$$

c) Si $\sigma^2(z_i) = 4z_i^2$, podemos obtener el estimador eficiente de MCG aplicando mínimos cuadrados ponderados con pesos

$$
\frac{1}{\sigma(z_i)} = \frac{1}{2z_i}
$$

lo que equivale a transformar el modelo dividiendo por $\sigma(z_i)$ y aplicar MCO al modelo transformado:

$$
\frac{y_i}{2z_i} = \alpha\frac{1}{2z_i} + \delta\frac{z_i}{2z_i} + u_i^*
$$

$$
\begin{aligned}
\begin{bmatrix} \widetilde{\alpha} \\ \widetilde{\delta} \end{bmatrix} &= \begin{bmatrix} \sum \frac{1}{z_i^2} & \sum \frac{1}{z_i} \\ \sum \frac{1}{z_i} & n \end{bmatrix}^{-1} \begin{bmatrix} \sum \frac{y_i}{z_i^2} \\ \sum \frac{y_i}{z_i} \end{bmatrix} = \begin{bmatrix} 5683 & 2058 \\ 2058 & 800 \end{bmatrix}^{-1} \begin{bmatrix} 18755 \\ 6835 \end{bmatrix} = \\
&= \frac{1}{311036} \begin{bmatrix} 800 & -2058 \\ -2058 & 5683 \end{bmatrix} \begin{bmatrix} 18755 \\ 6835 \end{bmatrix} = \\
&= \frac{1}{311036} \begin{bmatrix} 937570 \\ 245515 \end{bmatrix} = \\
&= \begin{bmatrix} 3.0143 \\ 0.7893 \end{bmatrix}
\end{aligned}
$$

$$
\begin{aligned}
var\left(\begin{bmatrix} \widetilde{\alpha} \\ \widetilde{\delta} \end{bmatrix}\right) &= \frac{1}{311036} \begin{bmatrix} 800 & -2058 \\ -2058 & 5683 \end{bmatrix} = \\
&= \begin{bmatrix} 0.00257 & \\ -0.0066 & 0.0183 \end{bmatrix}
\end{aligned}
$$

1.2 PROBLEMA 2

a)

$$\begin{aligned}
\widehat{\delta} &= \frac{\sum S_i y_i - \frac{1}{1000}\sum S_i \sum y_i}{\sum S_i^2 - \frac{1}{1000}\left(\sum S_i\right)^2} = \\
&= \frac{83312 - \frac{16707*4972}{1000}}{283539 - \frac{16707^2}{1000}} = \\
&= \frac{83312 - 83067.2}{283539 - 279123.8} = \\
&= \frac{244.8}{4415.2} = \\
&= 0.05544
\end{aligned}$$

$$\widehat{var}\left(\widehat{\delta}\right) = 0.27\frac{1}{4415.2}$$

Respecto a la interpretacion del resultado: "un año adicional de educación supone un incremento medio de ingresos salariales anuales de 55440 pesetas".

En cuanto al comentario sobre la consistencia: "al estar la variable explicativa medida con error, el error del modelo sera $v_i = u_i - \delta\varepsilon_i$, donde ε_i es el error de medida de la variable explicativa. Por tanto,la condicion $E(S_i v_i)$ será igual a $\delta E(\varepsilon_i)$, que no será en general igual a cero, de manera que el estimador MCO sera inconsistente".

b)

$$\begin{aligned}
\widehat{\delta}_{IV} &= \frac{\sum P_i y_i - \frac{1}{1000}\sum P_i \sum y_i}{\sum P_i S_i - \frac{1}{1000}\sum P_i \sum S_i} = \\
&= \frac{71404 - \frac{14343*4972}{1000}}{240466 - \frac{14343*16707}{1000}} = \\
&= \frac{71404 - 71313.4}{240466 - 239628.5} = \\
&= \frac{90.6}{837.5} = \\
&= 0.1082
\end{aligned}$$

$$\begin{aligned}
\widehat{var}\left(\widehat{\delta}_{IV}\right) &= \widehat{\sigma^2}\frac{\sum P_i^2 - \frac{1}{1000}\left(\sum P_i\right)^2}{\left[\sum P_i S_i - \frac{1}{1000}\sum P_i \sum S_i\right]^2} = \\
&= 0.27\frac{206469 - \frac{14343^2}{1000}}{\left[240466 - \frac{14343*16707}{1000}\right]^2} = \\
&= 0.27\frac{206469 - 205721.6}{\left[240466 - 239628.5\right]^2} = \\
&= 0.27\frac{747.4}{837.5^2}
\end{aligned}$$

$$s.e.\left(\widehat{\delta}_{IV}\right) = 0.017$$

$$H_0 : \delta = 0.12$$

$$t\left(\widehat{\delta}_{IV}\right)=\frac{0.1082-012}{0.017}=\frac{-0.0118}{0.017}$$

No rechazamos H_0.

c)

$$\begin{aligned}
H &= \frac{\left(\widehat{\delta}_{IV}-\widehat{\delta}\right)^2}{\widehat{var}\left(\widehat{\delta}_{IV}\right)-\widehat{var}\left(\widehat{\delta}\right)}=\frac{(0.1082-0.05544)^2}{0.27\left[\frac{747.4}{(837.2)^2}-\frac{1}{4415.2}\right]}= \\
&= \frac{0.05276^2}{0.27\,(0.001065-0.0002265)}= \\
&= 12.2888
\end{aligned}$$

$$H \underset{d}{\rightarrow} \chi_1^2$$

$$H>\chi^2_{1,1-\alpha} \qquad para\ \alpha=0.01$$

rechazamos $H_0:\delta_{OLS}=\delta_{IV}$ al 1%

$\Rightarrow S_i$ endógena
$\Rightarrow \widehat{\delta}$ inconsistente
Debemos utilizar $\widehat{\delta}_{IV}$.

1.3 PROBLEMA 3:

a) El modelo estimado es

$$\hat{Y}_i=\hat{a}+\hat{b}X_i+\hat{c}S_{1i}+\hat{d}F_{1i}.$$

Dadas las variables ficticias

$$\begin{aligned}
S_{1i} &= \begin{cases}1, & i \text{ es mujer,}\\ 0, & i \text{ es hombre,}\end{cases} \qquad S_{2i}=1-S_{1i}, \\
F_{1i} &= \begin{cases}1, & i \text{ con estudios,}\\ 0, & i \text{ sin estudios,}\end{cases} \qquad F_{2i}=1-F_{1i},
\end{aligned}$$

se tiene que,

$$\begin{array}{ll}
i=M+U: & \hat{Y}_i=\hat{a}+\hat{b}X_i+\hat{c}+\hat{d},\\
i=M+NU: & \hat{Y}_i=\hat{a}+\hat{b}X_i+\hat{c},\\
i=H+U: & \hat{Y}_i=\hat{a}+\hat{b}X_i+\hat{d},\\
i=H+NU: & \hat{Y}_i=\hat{a}+\hat{b}X_i.
\end{array} \tag{1}$$

(M: Mujer, U: Con estudios universitarios, H: Hombre, NU: Sin estudios universitarios).
* Modelo (i):

$$Y_i=\alpha+\beta X_i+\delta_1 S_{1i}+\delta_2 S_{2i}+\gamma F_{1i}+v_i.$$

No es posible estimar este modelo ante existencia de multicolinealidad, dado que

$$S_{1i}+S_{2i}=1, \quad \forall i.$$

* Modelo (ii):

$$Y_i=\beta X_i+\delta_1 S_{1i}+\delta_2 S_{2i}+\gamma_1 F_{1i}+\gamma_2 F_{2i}+w_i.$$

No es posible estimar este modelo ante existencia de multicolinealidad, dado que

$$S_{1i} + S_{2i} = F_{1i} + F_{2i}, \quad \forall i.$$

* Modelo (iii):

$$Y_i = \beta X_i + \delta_1 S_{1i} + \delta_2 S_{2i} + \gamma_1 F_{1i} + z_i,$$

de modo que si estimamos este modelo por MCO, se obtendrá que:

$$\begin{array}{ll} i = M + U: & \hat{Y}_i = \hat{\beta} X_i + \hat{\delta}_1 + \hat{\gamma}_1, \\ i = M + NU: & \hat{Y}_i = \hat{\beta} X_i + \hat{\delta}_1, \\ i = H + U: & \hat{Y}_i = \hat{\beta} X_i + \hat{\delta}_2 + \hat{\gamma}_1, \\ i = H + NU: & \hat{Y}_i = \hat{\beta} X_i + \hat{\delta}_2, \end{array}$$

y, por tanto, teniendo en cuenta (1), se obtiene que:

$$\begin{array}{rcl} \hat{\beta} & = & \hat{b}, \\ \hat{\delta}_1 & = & \hat{a} + \hat{c}, \\ \hat{\gamma}_1 & = & \hat{d}, \\ \hat{\delta}_2 & = & \hat{a}. \end{array}$$

* Modelo (iv):

$$\begin{array}{rcll} Y_i & = & \beta_1 X_i + \alpha_1 + \gamma F_{1i} + v_{1i}, & \text{si } i \text{ es hombre} \\ Y_i & = & \beta_2 X_i + \alpha_2 + \gamma F_{1i} + v_{2i}, & \text{si } i \text{ es mujer} \end{array}$$

y por tanto, si estimamos este modelo por MCO, se obtendrá que:

$$\begin{array}{ll} i = M + U: & \hat{Y}_i = \hat{\beta}_2 X_i + \hat{\alpha}_2 + \hat{\gamma}, \\ i = M + NU: & \hat{Y}_i = \hat{\beta}_2 X_i + \hat{\alpha}_2, \\ i = H + U: & \hat{Y}_i = \hat{\beta}_1 X_i + \hat{\alpha}_1 + \hat{\gamma}, \\ i = H + NU: & \hat{Y}_i = \hat{\beta}_1 X_i + \hat{\alpha}_1. \end{array}$$

Los estimadores de este último modelo no se pueden poner en términos de los correspondientes al modelo (1) dado que no son equivalentes. En el modelo (1) se permite un cambio de la ordenada en el origen en función del sexo y el nivel de estudios, pero sin embargo la pendiente es constante, a diferencia del modelo (iv), donde se permite también un cambio en la pendiente en función del sexo.

b) En este caso habría que modificar el modelo (1) a estimar, una posible forma sería:

$$Y_i = a + bX_i + cS_{1i} + dF_{1i} + eX_i F_{1i} + v_i,$$

de manera que para una persona con estudios universitarios la pendiente del modelo sería $b + e$ y para una sin estuidos sería igual a b.

c) Modelo a estimar:

$$Y_i = \alpha + \beta X_i + e_i, \tag{2}$$

Modelo verdadero:

$$Y_i = \alpha + \beta X_i + \delta S_{1i} + \gamma F_{i1} + \rho X_i F_{i1} + v_i. \tag{3}$$

Al estimar (2) siendo (3) el modelo verdadero nos encontramos ante un problema de omisión de variables relevantes, que supone obtener un estimador sesgado e inconsistente.