SOLUCIONES AL EXAMEN DE ECONOMETRÍA I

CURSO 2000/2001

PREGUNTA 1

Para estudiar si existen diferencias salariales entre hombres y mujeres en una industria, se toman 100 individuos al azar y se considera una función de regresión que relaciona a la vez el salario con los años de estudio y experiencia laboral para hombres y mujeres utilizando una variable ficticia. Las variables consideradas son:

$$\begin{aligned} W_i &= \text{salario hora del individuo } i \text{ en miles de pesetas.} \\ S_i &= \text{años de estudio del individuo } i. \\ E_i &= \text{años de experiencia laboral del individuo } i. \\ M_i &= \begin{cases} 1 \text{ si el individuo } i \text{ es mujer} \\ 0 \text{ si el individuo } i \text{ es hombre} \end{cases} \end{aligned}$$

La función de regresión considerada es:

$$W_i = \beta_1 + \beta_2 M_i + \beta_3 S_i + \beta_4 S_i \times M_i + \beta_5 E_i + \beta_6 E_i \times M_i + U_i,$$

donde U_i es un término de error que cumple todos los supuestos clásicos. Los valores de los coeficientes estimados por mínimos cuadrados ordinarios son:

$$\begin{aligned} \hat{\beta}_1 &= 2,0\ (0,10) \\ \hat{\beta}_2 &= -0,5\ (0,20) \\ \hat{\beta}_3 &= 0,09\ (0,01) \\ \hat{\beta}_4 &= -0,02\ (0,10) \\ \hat{\beta}_5 &= 0,1\ (0,05) \\ \hat{\beta}_6 &= -0,05\ (0,02) \end{aligned}$$

Entre paréntesis se encuentran los errores estándar de cada uno de los coeficientes estimados, que son computados suponiendo que se cumplen todos los supuestos clásicos. La suma de cuadrados de los residuos es $\widehat{\mathbb{U}}'\widehat{\mathbb{U}} = 240$. Por otro lado, se estima el modelo imponiendo la restricción $\beta_2 = \beta_4 = \beta_6 = 0$, obteniendo una suma de los cuadrados de los residuos de $\widetilde{\mathbb{U}}'\widetilde{\mathbb{U}} = 360$. Cuando proceda, escriba la hipótesis nula y alternativa e indique si el contrate ofrecido es exacto o asintótico, detallando, en el primer caso, los supuestos distribucionales que son necesarios.

1. (2 puntos) Interprete cada uno de los coeficientes del modelo y escriba las funciones de regresión estimadas para hombres y para mujeres.

2. (2 puntos) Contraste, al 5% de significación, que la variación de los salarios cuando el nivel de estudios varía en una unidad es idéntica para hombres y para mujeres.

3. (3 puntos) Contraste, al 5% de significación, que los modelos de salarios son idénticos para hombres y para mujeres .

4. (3 puntos) Explique el efecto que tendrá sobre los contrastes realizados que las varianzas de los términos de error, dados el nivel de estudios y experiencia laboral, para hombres y mujeres son diferentes. ¿Afectará esta circumstancia al sesgo de los estimadores propuestos? Comente sobre la validez de los contrastes propuestos realizados en los apartados 2 y 3 en este nuevo contexto. Justifique sus respuestas.

Solución:

1. Los dos modelos de regresión son los siguientes:

$$\begin{aligned} W_i &= \beta_1 + \beta_3 S_i + \beta_5 E_i + U_i \text{ si } i \text{ es hombre.} \\ W_i &= (\beta_1 + \beta_2) + (\beta_3 + \beta_4) S_i + (\beta_5 + \beta_6) E_i + U_i \text{ si } i \text{ es mujer.} \end{aligned}$$

   Por tanto, β_1 es la constante de la función de salarios de los hombres, interpretable como el salario medio que recibiría un hombre sin ningún año de estudios y sin experiencia profesional. β_2 se interpreta como la diferencia entre las contastes de la función de salarios de mujeres y hombres. β_3 se interpreta como la variación del salario de los hombres cuando aumenta en un año sus estudios, manteniendo constante su experiencia laboral. β_4 se interpreta como la diferencia entre la variacion en los salarios de mujeres y hombres cuando aumenta en un año sus estudios, manteniendo constante la experiencia laboral. β_5 se interpreta como la variación en los salarios de los hombres cuando aumenta en un año su experiencia laboral, manteniendo constante sus años de estudio. β_6 se interpreta como la diferencia entre la variacion en los salarios de mujeres y hombres cuando aumenta en un año su experiencia laboral, manteniendo los años de estudio constantes.

2. La hipótesis nula a contrastar es $H_0 : \beta_4 = 0$ versus $H_1 : \beta_4 \neq 0$. Se puede realizar el contraste

mediante un estadístico t para contratar la significatividad de la variable interactiva $S_i \times M_i$.

$$t = \frac{\hat{\beta}_4}{SE\left(\hat{\beta}_4\right)} = \frac{-0,02}{0,10} = -0,2.$$

Suponiendo que los errores son normales,

$$t \sim t_{94} \text{ bajo } H_0.$$

La aproximación asintótica establece que,

$$t \underset{aprox}{\sim} N(0,1) \text{ bajo } H_0.$$

De hecho, dado el elevado número de observaciones disponibles, $t_{94} \approx N(0,1)$ y los valores críticos en los dos contrastes son prácticamente idénticos. El valor crítico, al 5% de significación, es $1,985$ utilizando la t_{94} y $1,96$ utilizando la $N(0,1)$. $|t|$ es muy inferior al valor crítico, por lo que no podemos rechazar la hipótesis nula.

3. La hipótesis a contrastar es $H_0 : \beta_2 = \beta_4 = \beta_6 = 0$. Con la información disponible, solo podemos realizar un contraste tipo F basado en la suma de los cuadrados de los residuos. El estadístico del contraste es:

$$F = \frac{\left(\widetilde{\mathbb{U}}'\widetilde{\mathbb{U}} - \widehat{\mathbb{U}}'\widehat{\mathbb{U}}\right)/q}{\widehat{\mathbb{U}}'\widehat{\mathbb{U}}/(n-k)},$$

donde q es el número de restricciones impuestas, $q = 3$, y k el número de parámetros del modelo no restringido, $k = 6$. Por tanto,

$$F = \frac{(360 - 240)/3}{360/(100-6)} = 10,444.$$

Suponiendo que la distribución condicional de los errores, dadas las variables explicativas, es normal.

$$F \sim F_{(3,94)} \text{ bajo } H_0.$$

Sin realizar el supuesto de normalidad, sabemos que,

$$F \underset{aprox}{\sim} \chi^2_3/3 \text{ bajo } H_0.$$

De hecho, dado el elevado número de observaciones disponibles, $3F_{(3,94)} \approx \chi^2_3$ y los valores críticos en los dos contrastes son prácticamente idénticos. El valor crítico, al 5% de significación, es $2,715$ utilizando la $F_{(3,94)}$ y $7,82/3 = 2,606\,7$ utilizando la aproximación asintótica (obsérvese que los valores críticos computados por cualquiera de los dos procedimientos son practicamente idénticos). Por tanto, rechazamos la hipótesis nula $(10,444 > 2,715 > 2,6067)$.

4. Se nos plantea un problema de heteroskedasticidad. En particular, pudiera ocurrir que,
$$E\left(U_i^2\middle|\, E_i, S_i, M_i\right) = \sigma_1^2 M_i + \sigma_2^2\left(1 - M_i\right) = \begin{cases} \sigma_1^2 \text{ si } i \text{ es mujer} \\ \sigma_2^2 \text{ si } i \text{ es hombre} \end{cases},\ \sigma_1^2 \neq \sigma_2^2.$$
El enunciado no descarta que σ_1^2 y σ_2^2 no fueran constantes, sino funciones de E_i y S_i. El enunciado del problema dice que los errores estándar han sido computados suponiendo que los supuestos clásicos se cumplen, entre los que se encuentra que las varianzas condicionales de los errores son constantes, por tanto, estos errores estándar no aproximan de forma adecuada las desviaciones típicas de los estimadores y las conclusiones basadas en el estadístico t de la pregunta 2 pueden ser incorrectas, no pueden justificarse desde un punto de vista científico. La distribución del estadístico F bajo la nula, basado en la suma de los cuadrados de los residuos, tanto en su versión exacta, suponiendo normalidad, como en su versión asintótica, depende del supuesto de homoskedasticidad. Por tanto, las conclusiones a las que llegamos en la pregunta 3 pueden ser incorrectas, no tienen justificación desde un punto de vista científico. Sin embargo, los estimadores MCO siguen siendo insesgados.

PROBLEMA 2

En un estudio de mercado realizado por una empresa se ha considerado el siguiente modelo de regresión, que relaciona ventas con gasto en publicidad:

$$Y_i = \beta_1 + \beta_2 X_{i2} + V_i,\ i = 1, ..., n \tag{1}$$

donde X_{i2} e Y_i representan, respectivamente, el gasto en publicidad (en miles de pesetas) y las ventas (en millones de pesetas) de la empresa i, V_i es un término de error. El modelo se ha estimado en base a una muestra aleatoria de 32 empresas, resultando el modelo estimado:

$$\hat{Y}_i = 2,554 + 0,327 X_{i2},\ i = 1, ..., n \tag{2}$$

siendo el coeficiente de deperminación $\tilde{R}^2 = 0,122$. Un analista de mercado que participa en el estudio, ha sugerido que los gastos en investigación y desarrollo (I+D) y el precio del bien pueden tener también capacidad de explicar el comportamiento de las ventas, proponiendo el siguiente modelo alternativo:

$$Y_i = \beta_1 + \beta_2 X_{i2} + \beta_3 X_{i3} + \beta_4 X_{i4} + U_i,\ i = 1, ..., n \tag{3}$$

donde X_{i3} representa los gastos anuales en I+D, en miles de pesetas, y X_{i4} representa el precio del bien (en pesetas) fijado por el principal competidor de la empresa, U_i es un término de error. Esta segunda ecuación se estima con los datos disponibles, dando lugar a un coeficiente de determinación $\hat{R}^2 = 0,216$. Suponemos que $(Y_1, X_{12}, X_{13}, X_{14}), ..., (Y_n, X_{n2}, X_{n3}, X_{n4})$ son observaciones independientes e idénticamente distribuidas como (Y, X_2, X_3, X_4). Cuando proceda, escriba la hipótesis nula y alternativa, suponga que bajo la hipótesis nula se cumplen todos los supuestos clásicos, e indique si el contrate ofrecido es exacto o asintótico, detallando, en el primer caso, los supuestos distribucionales que son necesarios.

1. (2 puntos) Suponiendo que la esperanza condicional de Y dado X es lineal, y por lo tanto idéntica al predictor lineal óptimo, exprese los coeficientes β_1 y β_2 en términos de esperanzas, varianzas y covarianzas.

2. (2 puntos) Suponiendo que los errores U_i cumplen todos los supuestos clásicos. Contraste al 5% de significación la significatividad del gasto en publicidad en el modelo representado por la Ecuación (1).

3. (3 puntos) Contraste al 5% de significación si es necesaria la inclusión de las variables explicativas X_{i3} y X_{i4} en el modelo representado por la Ecuación (1). Esto es, contraste si X_{i3} y X_{i4} son significativas de forma conjunta. Justifique la respuesta.

4. (3 puntos) Sobre la base de los contrastes realizados, discuta la insesgadez y eficiencia de los estimadores MCO del parámetro β_2 obtenidos en los dos apartados anteriores.

Solución

1. El predictor lineal óptimo de Y dado X lo podemos expresar como,

$$E^*(Y|X_2) = \beta_1 + \beta_2 X_2,$$

donde,

$$\begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} = \arg\min_{a,b} E\left[(Y - a - bX_2)^2\right].$$

Las ecuaciones normales son,

$$
\begin{aligned}
-2E\left(Y-\beta_1-\beta_2 X_2\right) &= 0 \\
-2E\left[\left(Y-\beta_1-\beta_2 X_2\right) X_2\right] &= 0,
\end{aligned}
$$

cuya solución da lugar a,

$$
\begin{aligned}
\beta_1 &= E\left(Y\right)-\beta E\left(X_2\right), \\
\beta_2 &= \frac{Cov\left(Y, X_2\right)}{Var\left(X_2\right)},
\end{aligned}
$$

por tanto,

$$
E^*\left(Y \mid X_2\right)=E\left(Y\right)+\frac{Cov\left(Y, X_2\right)}{Var\left(X_2\right)}\left[X_2-E\left(X_2\right)\right].
$$

2. La hipótesis nula y alternativa se pueden escribir como,

$$
H_0: \beta_2=0 \quad vs \quad H_1: \beta_2 \neq 0 \ .
$$

La información disponible solo permite realizar el contraste utilizando un estadístico F, expresado en términos del coeficiente de determinación. Observesé que la hipótesis nula indica que ninguna variable explicativa del modelo es significativa, por tanto el estadístico del contraste de significación global es:

$$
F=\frac{\tilde{R}^2/\left(k-1\right)}{\left(1-\tilde{R}^2\right)/\left(n-k\right)}=\frac{0,122}{(1-0,122)/30}=4,16
$$

teniendo en cuenta que en nuestro caso $n=32$ y $k=2$. Suponiendo que los errores son condicionalmente normales,

$$
F \sim F_{(1,30)} \text{ bajo } H_0.
$$

La aproximación asintótica establece que,

$$
F \underset{aprox}{\sim} \chi^2_{(1)} \text{ bajo } H_0.
$$

El valor crítico al 5% de significación es aproximadamente $7,5625$ utilizando la $F_{(1,30)}$, y $6,649$ utilizando la $\chi^2_{(1)}$. El valor del estadístico F es mayor el valor crítico utilizando cualquiera de los dos criterios, y, por tanto, la hipótesis nula se rechaza al nivel de significación propuesto.

3. La hipótesis a contrastar es la siguiente,

$$H_0 : \beta_3 = \beta_4 = 0 \quad vs \quad H_1 : \beta_3 \neq 0 \text{ó } \beta_4 \neq 0$$

La información disponible obliga a utilizar un estadístico F en términos de los coeficientes de determinación del modelo restringido y no restringido. Esto es,

$$F = \frac{\left(\hat{R}^2 - \tilde{R}^2\right)/q}{(1-\hat{R}^2)/(n-k)} = \frac{(0,216 - 0,122)/2}{(1-0,216)/28} = 1,99,$$

teniendo en cuenta que en nuestro caso $q = 2$, $n = 32$ y $k = 4$. Suponiendo que los errores son condicionalmente normales,

$$F \sim F_{(2,30)} \text{ bajo } H_0.$$

La aproximación asintótica establece que,

$$F \underset{aprox}{\sim} \chi^2_{(2)}/2 \text{ bajo } H_0.$$

El valor crítico utilizando la $F_{(2,30)}$ es $3,3158 > 1,99$, por lo que se no se puede rechazar H_0. El valor crítico utilizando la aproximación asintótica es $5,99147/2 = 2,9957 > 1,99$ $\left(\Pr\left(\chi^2_{(2)} > 5,99147\right) = 0,05\right)$, por lo que tampoco se rechaza la hipótesis nula utilizando esta aproximación. Obsérvese que el contraste exacto es más conservador que el asintótico, al tener una probabilidad del tipo I de error menor.

4. Las variables X_3 y X_4 no son significativas de forma conjunta, por lo que se espera que el estimador MCO de β_2 en modelo restringido (1) sea más eficiente que el estimador en el modelo no restringido (3).

PROBLEMA 3

En una industria, las empresas que la componen relacionan sus stocks *esperados* de productos acabados, Y^e, con sus ventas anuales *esperadas*, X^e, de acuerdo con el modelo lineal:

$$Y^e = \alpha + \beta X^e + \varepsilon,$$

donde α y β son coeficientes desconocidos y ε es un término de error con media cero e incorrelacionado con X^e. Las ventas anuales *observadas*, X,difieren de las ventas esperadas por un error de medida, U, de media cero e independiente de X^e y ε, esto es,

$$X = X^e + U, \qquad ((1))$$

Por otro lado, los stocks de productos acabados *observados,* Y, difieren de los experados por un error de medida, V, con media cero e independiente de X^e, Y^e y de ε,

$$Y = Y^e - V,$$

1. Discuta de forma razonada las propiedades de consistencia del estimador de mínimos cuadrados de β, basado en observaciones $(Y_1, X_1), ..., (Y_n, X_n)$ independientes e idénticamente distribuidas como (Y, X) :

$$\hat{\beta} = \frac{\sum_{i=1}^{n} Y_i \left(X_i - \bar{X}\right)}{\sum_{i=1}^{n} \left(X_i - \bar{X}\right)^2},$$

en cada uno de los siguientes escenarios:

(a) (2 puntos) Cuando $E\left(V^2\right) = 0$, esto es, Y_i^e son observadas, pero $E\left(U^2\right) \neq 0$.

(b) (2 puntos) Cuando $E\left(U^2\right) = 0$, esto es, X_i^e son observadas, $E\left(V^2\right) \neq 0$.

(c) (2 puntos) Cuando $E(UV) = E\left(U^2\right) \neq 0$, lo que implica que $U = V$ con probabilidad 1, y que el coeficiente de correlación entre U y V es igual a 1.

2. (4 puntos) La teoría nos dice que la cantidad de trabajo, L, empleada por la empresa, es una función lineal determinística de las ventas anuales esperadas, es decir,

$$L = \theta_1 + \theta_2 X^e \text{ con probabilidad } 1,$$

donde θ_1 y θ_2 son constantes desconocidas. Esto implica que el coeficiente de correlación entre L y X^e es igual a 1 cuando $\theta_2 > 0$, e igual a -1 cuando $\theta_2 < 0$. Proponga un vector de instrumentos, y demuestre que es válido para estimar de forma consistente el vector de parámetros (α, β). Proporcione una expresión para el estimador de variables instrumentales basado en observaciones $(Y_1, X_1, L_1), ..., (Y_n, X_n, L_n)$ independientes e identicamente distribuidas como (Y, X, L).

Solución

1. Escribimos,

$$Y_i = \alpha + \beta X_i + (\varepsilon_i - \beta U_i + V_i)$$

$$\hat{\beta} = \beta + \frac{\frac{1}{n}\sum_{i=1}^{n}\left(\varepsilon_i - \beta U_i + V_i\right)\left(X_i - \bar{X}\right)}{\frac{1}{n}\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2},$$

Por la Ley de los Grandes Números, las medias muestrales convergen en probabilidad a sus análogos poblacionales, por lo que tenemos que,

$$\begin{aligned}
&\bar{X} \xrightarrow{p} E\left(X\right) \\
&\frac{1}{n}\sum_{i=1}^{n}\left(\varepsilon_i - \beta U_i + V_i\right) \xrightarrow{p} E\left[\varepsilon - \beta U + V\right] \\
&\frac{1}{n}\sum_{i=1}^{n}\left(\varepsilon_i - \beta U_i + V_i\right) X_i \xrightarrow{p} E\left[\left(\varepsilon - \beta U + V\right) X\right] \\
&\frac{1}{n}\sum_{i=1}^{n} X_i^2 \xrightarrow{p} E\left(X^2\right)
\end{aligned}$$

Los supuestos sobre los que descansa nuestro modelo implican que,

$$\begin{aligned}
E\left(X\right) &= E\left(X^e\right) \\
E\left[\varepsilon - \beta U + V\right] &= 0 \\
E\left[\left(\varepsilon - \beta U + V\right) X\right] &= E\left[\left(\varepsilon - \beta U + V\right)\left(X^e + U\right)\right] = -\beta E\left(U^2\right) + E\left(UV\right) \\
E\left(X^2\right) &= E\left(X^{e2}\right) + E\left(U^2\right).
\end{aligned}$$

Por tanto,

$$\begin{aligned}
\frac{1}{n}\sum_{i=1}^{n}\left(\varepsilon_i - \beta U_i + V_i\right)\left(X_i - \bar{X}\right) &= \frac{1}{n}\sum_{i=1}^{n}\left(\varepsilon_i - \beta U_i + V_i\right) X_i - \bar{X}\frac{1}{n}\sum_{i=1}^{n}\left(\varepsilon_i - \beta U_i + V_i\right) \\
&\xrightarrow{p} E\left[\left(\varepsilon - \beta U + V\right) X\right] - E\left(X^e\right) \times 0 \\
&= -\beta E\left(U^2\right) + E\left(UV\right).
\end{aligned}$$

y

$$\frac{1}{n}\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2 = \frac{1}{n}\sum_{i=1}^{n} X_i{}^2 - \bar{X}^2 \xrightarrow{p} E\left(X^2\right) - E\left(X\right)^2 = E\left(X^{e2}\right) + E\left(U^2\right) - E\left(X^e\right)^2.$$

Por lo tanto,

$$\hat{\beta} \xrightarrow{p} \beta + \frac{-\beta E\left(U^2\right) + E\left(UV\right)}{E\left(X^{e2}\right) + E\left(U^2\right) - E\left(X^e\right)^2} = \frac{-\beta E\left(U^2\right) + E\left(UV\right)}{Var\left(X^e\right) + E\left(U^2\right)},$$

observando que $Var\left(X^e\right) = E\left(X^{e2}\right) - E\left(X^e\right)^2$. A partir de esta expresión tenemos que:

(a)

$$\hat{\beta} \xrightarrow{p} \beta + \frac{-\beta E\left(U^2\right) + E\left(UV\right)}{Var\left(X^e\right) + E\left(U^2\right)} = \beta\left[1 - \frac{E\left(U^2\right)}{Var\left(X^e\right) + E\left(U^2\right)}\right] < \beta,$$

teniendo en cuenta que $E\left(UV\right) = 0$, al ser V una contante igual a cero. El estimador es inconsistente.

(b)

$$\hat{\beta} \xrightarrow{p} \beta + \frac{-\beta E\left(U^2\right) + E\left(UV\right)}{Var\left(X^e\right) + E\left(U^2\right)} = \beta,$$

teniendo en cuenta que se ha supuesto que U es una constante igual a cero, y por tanto $E\left(U^2\right) = 0$ y $E\left(UV\right) = 0$.

(c)

$$\hat{\beta} \xrightarrow{p} \beta + \frac{-\beta E\left(U^2\right) + E\left(UV\right)}{Var\left(X^e\right) + E\left(U^2\right)} = \beta + \frac{\left(1-\beta\right) E\left(U^2\right)}{Var\left(X^e\right) + E\left(U^2\right)} \begin{cases} = \beta \text{ si } \beta = 1 \\ \neq \beta \text{ si } \beta \neq 1 \end{cases},$$

teniendo en cuenta que $E\left(UV\right) = E\left(U^2\right)$. El estimador es consistente solamente cuando el stock esperado aumenta en la misma proporción que las ventas esperadas, esto es, cuando $\beta = 1$.

2. El vector de instrumentos a utilizar es

$$W = \begin{pmatrix} 1 \\ L \end{pmatrix},$$

esto es se instrumentaliza la variable X que está correlacionada con el término de error ε. La primera condición que ha de cumplirse es,

$$E\left(W\varepsilon\right) = 0, \tag{4}$$

Por un lado,

$$E\left(1 \times \varepsilon\right) = E\left(\varepsilon\right) = 0,$$

por otro,

$$E\left(L\varepsilon\right) = E\left[\left(\theta_1 + \theta_2 X^e\right)\varepsilon\right] = \theta_1 E\left(\varepsilon\right) + \theta_2 E\left(X^e \varepsilon\right) = 0.$$

Por lo que la condición (4) se cumple. Las dos ecuaciones en (4) pueden escribirse como:

$$\begin{aligned} E\left(Y - \alpha - \beta X\right) &= 0 \\ E\left[\left(Y - \alpha - \beta X\right) L\right] &= 0, \end{aligned}$$

o de forma más compacta como:

$$\begin{pmatrix} E(Y) \\ E(YL) \end{pmatrix} - \begin{bmatrix} 1 & E(X) \\ E(L) & E(XL) \end{bmatrix} (\alpha, \beta) = 0, \tag{5}$$

ecuación que solo tiene solución en términos de (α, β) cuando,

$$\begin{vmatrix} 1 & E(X) \\ E(L) & E(XL) \end{vmatrix} = Cov(X, L) \neq 0,$$

que es la segunda condición que ha de cumplir nuestro vector de instrumentos. Observamos que:

$$\begin{aligned} Cov(X, L) &= E(XL) - E(X) E(L) \\ &= E[(X^e + U)(\theta_1 + \theta_2 X^e)] - E(X^e + U) E(\theta_1 + \theta_2 X^e) \\ &= \theta_2 E\left(X^{e2}\right) + \theta_1 E(X^e) - \theta_1 E(X^e) - \theta_2 E(X^e)^2 \\ &= \theta_2 Var(X^e) \neq 0 \text{ siempre que } \theta_2 \neq 0. \end{aligned}$$

Concluimos que nuestro vector de instrumentos es válido. La solución del sistema de ecuaciones (5) nos permite expresar los parámetros del modelo en términos de momentos. Esto es,

$$\begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{bmatrix} 1 & E(X) \\ E(L) & E(XL) \end{bmatrix}^{-1} \begin{pmatrix} E(Y) \\ E(YL) \end{pmatrix} = \begin{pmatrix} E(Y) - \frac{Cov(Y,L)}{Cov(X,L)} E(X) \\ \frac{Cov(Y,L)}{Cov(X,L)} \end{pmatrix}.$$

Dadas observaciones $(Y_1, X_1, L_1), ..., (Y_n, X_n, L_n)$ independientes e identicamente distribuidas como (Y, X, L). Los parámetros se estiman por su análogo muestral, substituyendo medias muestrales por sus análogos poblacionales. Por tanto el estimador de variables instrumentales es:

$$\begin{aligned} \begin{pmatrix} \hat{\alpha}^{VI} \\ \hat{\beta}^{VI} \end{pmatrix} &= \begin{bmatrix} 1 & \frac{1}{n}\sum_{i=1}^n X_i \\ \frac{1}{n}\sum_{i=1}^n L_i & \frac{1}{n}\sum_{i=1}^n X_i L_i \end{bmatrix}^{-1} \begin{pmatrix} \frac{1}{n}\sum_{i=1}^n Y_i \\ \frac{1}{n}\sum_{i=1}^n X_i L_i \end{pmatrix} \\ &= \left\{ \sum_{i=1}^n \begin{bmatrix} 1 & X_i \\ L_i & X_i L_i \end{bmatrix} \right\}^{-1} \sum_{i=1}^n \begin{pmatrix} Y_i \\ X_i L_i \end{pmatrix}, \end{aligned}$$

lo que podemos expresar también como:

$$\begin{aligned} \hat{\alpha}^{VI} &= \bar{Y} - \hat{\beta}^{VI} \bar{X} \\ \hat{\beta}^{VI} &= \frac{\sum_{i=1}^n \left(L_i - \bar{L}\right)\left(Y_i - \bar{Y}\right)}{\sum_{i=1}^n \left(L_i - \bar{L}\right)^2}, \end{aligned}$$

donde $\bar{Y} = n^{-1}\sum_{i=1}^n Y_i$, $\bar{X} = n^{-1}\sum_{i=1}^n X_i$ y $\bar{L} = n^{-1}\sum_{i=1}^n L_i$.