

## AYUDANTIA N°4 - ECONOMETRÍA

**Profesor:** Juan Urquiza

**Ayudante:** Valentina Andrade (vandrade@uc.cl)

---

### TEMA I

Suponga que la productividad promedio de los trabajadores de las empresas manufactureras (*avgprod*) depende, entre otras cosas, del promedio de horas de capacitación (*avgtrain*) y de la habilidad promedio de los trabajadores (*avghabil*) tal que:

$$avgprod = \beta_0 + \beta_1 \times avgtrain + \beta_2 \times avghabil + u,$$

donde se satisfacen todos los supuestos revisados en clase.

Interesado en estudiar la relación entre productividad y horas de capacitación, un investigador decide estimar la regresión simple de *avgprod* sobre *avgtrain*, debido a que no dispone de datos sobre la habilidad de los trabajadores. Si usted supiera que el gobierno ha otorgado subvenciones a las empresas cuyos trabajadores tienen habilidades inferiores al promedio, ¿cuál sería el sesgo probable en el estimador de  $\beta_1$  que se obtiene a partir de la regresión simple de *avgprod* sobre *avgtrain*? Discuta.

### TEMA II

Usted desea estimar la relación entre desempeño académico e ingreso familiar. Para ello, consigue datos sobre las calificaciones promedio en la universidad (*colGPA*), las calificaciones promedio en el bachillerato (*hsGPA*) y el puntaje en la prueba de admisión universitaria (*SAT*). Entonces, para probar si el ingreso familiar tiene relación con el desempeño en la universidad, postula el siguiente modelo:

$$colGPA = \beta_0 + \beta_1 \times faminc^* + \beta_2 \times hsGPA + \beta_3 \times SAT + u,$$

donde *faminc\** es el ingreso familiar anual.

Sin embargo, el ingreso familiar no está disponible y entonces debe recurrir al ingreso familiar (*faminc*) reportado por los propios estudiantes. Suponiendo que se cumple el supuesto de errores clásicos en variables, discuta las consecuencias de utilizar el ingreso familiar reportado en lugar del ingreso familiar verdadero.

### TEMA III

Un estudio reciente relaciona el promedio de horas de televisión semanales ( $tvhours^*$ ) que ven los niños con su edad ( $age$ ), los años de educación de la madre ( $motheduc$ ), los años de educación del padre ( $fatheduc$ ) y el número de hermanos ( $sibs$ ):

$$tvhours^* = \beta_0 + \beta_1 \times age + \beta_2 \times age^2 + \beta_3 \times motheduc + \beta_4 \times fatheduc + u.$$

A usted le preocupan las conclusiones del estudio ya que en el mismo no se dispone de la variable  $tvhours^*$ , sino que se cuenta con las horas semanales de televisión ( $tvhours$ ) reportadas por alguno de los padres.

- ¿Qué condiciones se requieren para que en esta aplicación se cumpla con el supuesto de errores clásicos en las variables?
- ¿Considera usted que es probable que se cumplan dichas condiciones en este contexto? Explique.

### TEMA IV (TAREA 4)

Considere los resultados del siguiente estudio sobre la “brecha de género” en los Estados Unidos, donde se compara la compensación salarial de altos ejecutivos de firmas que cotizan en la bolsa. En particular, se dispone de los salarios anuales (en miles de dólares) de 46,000 altos ejecutivos, el valor de mercado de la firma ( $mktvalue$ , en millones de dólares) como una medida del tamaño de la firma, y una variable binaria ( $mujer$ ) que toma el valor de 1 para mujeres y 0 para hombres.

Los resultados de la estimación por Mínimos Cuadrados Ordinarios (MCO) de distintos modelos se presentan a continuación:

$$(M1): \quad \log(\widehat{\text{salar}}) = 6.48 - 0.44 \times mujer \\ (0.01) \quad (0.05)$$

$$(M2): \quad \log(\widehat{\text{salar}}) = 3.68 - 0.28 \times mujer + 0.37 \times \log(mktvalue) \\ (0.03) \quad (0.04) \quad (0.004)$$

donde los errores estándar se reportan entre paréntesis.

- Interprete económicamente los coeficientes del modelo (M2) y, en base a la información disponible, discuta si es más probable que las firmas grandes tengan altos ejecutivos mujeres que las firmas pequeñas. Justifique su respuesta.

- b. El estudio también reporta los resultados de una especificación que incorporara como variable de control adicional al retorno de la acción de la firma (*stockreturn*, en porcentaje). Refiérase a esta especificación como modelo (M3), y explique las consecuencias de incorporar a dicha variable sobre las propiedades de los estimadores de MCO, tanto si la considera relevante como si la considera irrelevante.
- c. Reflexionando acerca de los resultados del estudio, se preocupa por eventuales errores de medición en los salarios, dado que los mismos fueron reportados por los propios ejecutivos. Suponiendo un error de medición multiplicativo tal que  $\text{salario} = (a_0 \times \text{salario}^*)$ , donde  $a_0 > 0$ , de modo tal que el error de medición  $e_0 = \log(a_0)$ , explique cómo este error de medición podría afectar a la estimación por MCO.

## TEMA I

Suponga que la productividad promedio de los trabajadores de las empresas manufactureras ( $avgprod$ ) depende, entre otras cosas, del promedio de horas de capacitación ( $avgtrain$ ) y de la habilidad promedio de los trabajadores ( $avghabil$ ) tal que:

$$avgprod = \beta_0 + \beta_1 \times avgtrain + \beta_2 \times avghabil + u, \quad \text{modelo original}$$

donde se satisfacen todos los supuestos revisados en clase.

Interesado en estudiar la relación entre productividad y horas de capacitación, un investigador decide estimar la regresión simple de  $avgprod$  sobre  $avgtrain$ , debido a que no dispone de datos sobre la habilidad de los trabajadores. [Si usted supiera que el gobierno ha otorgado subvenciones a las empresas cuyos trabajadores tienen habilidades inferiores al promedio, ¿cuál sería el sesgo probable en el estimador de  $\beta_1$  que se obtiene a partir de la regresión simple de  $avgprod$  sobre  $avgtrain$ ? Discuta.]

- Identificar el problema.

→ Sesgo de variable omitida.

- Puedo tener una hipótesis de como es  $\beta_2$   
 $\beta_2 > 0$  (efecto positivo sobre la productividad)

- ¿Cómo se relaciona entre capacitación y productividad?  
 $Cov(\text{capacitación}, \text{habilidades}) > 0$

¡Ahora vamos a mirar lo que nos dice la teoría..!

$$Avgprod = \beta_0 + \tilde{\beta}_1 \cdot avgtrain + u$$

$$\tilde{\beta}_1 = \beta_1 + \underbrace{\beta_2 \cdot avghabil}$$

$$\tilde{\beta}_1 < \beta_1, \quad \text{correlación } avgtrain \text{ y } avghabil \text{ (+)}$$



## TEMA II

Usted desea estimar la relación entre desempeño académico e ingreso familiar. Para ello, consigue datos sobre las calificaciones promedio en la universidad (colGPA), las calificaciones promedio en el bachillerato (hsGPA) y el puntaje en la prueba de admisión universitaria (SAT). Entonces, para probar si el ingreso familiar tiene relación con el desempeño en la universidad, postula el siguiente modelo:

$$\text{colGPA} = \beta_0 + \beta_1 \times \text{faminc}^* + \beta_2 \times \text{hsGPA} + \beta_3 \times \text{SAT} + u, \quad \text{modelo original}$$

donde  $\text{faminc}^*$  es el ingreso familiar anual.

Sin embargo, el ingreso familiar no está disponible y entonces debe recurrir al ingreso familiar ( $\text{faminc}$ ) reportado por los propios estudiantes. Suponiendo que se cumple el supuesto de errores clásicos en variables, discuta las consecuencias de utilizar el ingreso familiar reportado en lugar del ingreso familiar verdadero.

↳ Error de medición en la variable independiente (x)

$$\text{faminc} = \text{faminc}^* + e_1$$

$$\text{colGPA} = \beta_0 + \beta_1 \cdot \text{faminc} + \beta_2 \cdot \text{hsGPA} + \beta_3 \text{SAT} + u + \underbrace{\beta_1 \cdot e_1}_{\text{error compuesto.}}$$

### 1. Propiedades MCO

Sesgo : sesgo de atenuación

Consistencia

$$\hat{\beta}_1 = \beta_1 \cdot \frac{\text{Var}(\text{faminc}^*)}{\text{Var}(\text{faminc})}$$

↓

$$\text{Var}(\text{faminc}^*) + \underbrace{\text{Var}(e_1)}_{\text{becas}}$$

$\hat{\beta}_1 \rightarrow \text{sesgo}$

### 2. ¿Tiene esto un efecto en inferencia?

$\hat{\beta}_1$  si es muy grande el sesgo de atenuación, que  $\hat{\beta}_1 \approx 0$

$$H_0: \tilde{\beta}_1 = 0$$

$$H_1: \hat{\beta}_1 > 0$$

} Podría ocurrir que no pueda rechazar  $H_0$ .

### TEMA III - Error medición en la variable dependiente (4)

Un estudio reciente relaciona el promedio de horas de televisión semanales ( $tvhours^*$ ) que ven los niños con su edad ( $age$ ), los años de educación de la madre ( $motheduc$ ), los años de educación del padre ( $fatheduc$ ) y el número de hermanos ( $sibs$ ):

$$tvhours^* = \beta_0 + \beta_1 \times age + \beta_2 \times age^2 + \beta_3 \times motheduc + \beta_4 \times fatheduc + u.$$

A usted le preocupan las conclusiones del estudio ya que en el mismo no se dispone de la variable  $tvhours^*$ , sino que se cuenta con las horas semanales de televisión ( $tvhours$ ) reportadas por alguno de los padres.

- ¿Qué condiciones se requieren para que en esta aplicación se cumpla con el supuesto de errores clásicos en las variables?
- ¿Considera usted que es probable que se cumplan dichas condiciones en este contexto? Explique.

a. Para que el error clásico de variables se cumpla tenemos

$$tvhours = tvhours^* + e_0$$

$$E(u) = 0$$
$$Var(u|x) = 0$$

$$\left\{ \begin{array}{l} \bullet E(e_0) = 0 \\ \bullet Cov(tvhours^*, e_0) = 0 \end{array} \right.$$

→ y para cualquier otra combinación con  $x_i$

Con esto el estimador de MCO es insesgado y consistente

b. En este ejemplo, los errores clásicos de variable son difíciles de cumplir.

► Para niños que no ven TV,  $tvhours^* = 0$ , y es muy posible que reporten cero horas viendo TV

⇒ Luego, si  $tvhours^* = 0$ , luego  $e_0 = 0$  con alta probabilidad  
si  $tvhours^* > 0$ , el error puede ser positivo o negativo

pero dado que  $tvhours \geq 0$ ,  $e_0$  satisface que  
$$e_0 \geq -tvhours^*$$

Luego,  $e_0$  y  $tvhours^*$  están posiblemente correlacionadas

▷ Padres con más años de educación