Proyecto de Innovación Docente: Elaboración de materiales interactivos y multidisciplinares para favorecer el aprendizaje y evaluación en los estudios de grado y postgrado

Financiado por el Vicerrectorado de Formación Permanente, Transformación Docente y Ocupación de la Universitat de València
PROYECTO DE INNOVACION EDUCATIVA CONSOLIDADO-PIEC, UV-SFPIE_PIEC-2736595

# Propiedades de los estimadores MCO

Víctor Enrique Barrios Pérez

Jorge Belaire Franch

$$CO = \beta_0 + \beta_1 REN + u$$

*Función de regresión poblacional*

$$E(CO|REN) = \beta_0 + \beta_1 REN \text{ si } E(u|REN) = E(u) = 0$$

$$u = CO - E(CO|REN) \qquad \beta_1 = \Delta E(CO|REN) / \Delta REN$$

*Función de regresión muestral*

$$CO_i = \hat{\beta}_0 + \hat{\beta}_1 REN_i$$

$$\hat{u}_i = CO_i - CO_i = CO_i - \hat{\beta}_0 - \hat{\beta}_1 REN_i$$

•**Los estimadores por Mínimos Cuadrados Ordinarios son aquellos que resuelven en siguiente problema de minimización**:

$$\min_{\hat{\beta}_0,\hat{\beta}_1} S = \min_{\hat{\beta}_0,\hat{\beta}_1} \sum_{i=1}^{n} \hat{u}^2 = \sum_{i=1}^{n} \left( CO_i - \hat{\beta}_0 - \hat{\beta}_1 REN_i \right)^2$$

- **Las condiciones de primer orden** $\dfrac{\partial S}{\partial \hat{\beta}_0} = 0$ **y** $\dfrac{\partial S}{\partial \hat{\beta}_1} = 0$ son :

$$\overline{CO} = \hat{\beta}_0 + \hat{\beta}_1 \overline{REN} \rightarrow \boxed{\hat{\beta}_0 = \overline{CO} - \hat{\beta}_1 \overline{REN}}$$

$$\boxed{\hat{\beta}_1 = \frac{\sum_{i=1}^{n} \left( REN_i - \overline{REN} \right)\left( CO_i - \overline{CO} \right)}{\sum_{i=1}^{n} \left( REN_i - \overline{REN} \right)^2} = \frac{\operatorname{cov}(REN_i, CO_i)}{\operatorname{var}(REN_i)}}$$

# Hipótesis estadísticas

Bajo los supuestos S.1 a S.4:

- S.1 El modelo poblacional es **lineal** en los parámetros.
- S.2 Utilizamos una muestra **aleatoria** de tamaño n.
- S.3 Media condicional nula $E(u_i|x_{i1},\ldots,x_{ik}) = E(u) = 0$.

$\rightarrow$ *Los factores incluidos en u no están correlacionados con los regresores del modelo.*

- S.4 No hay multicolinealidad perfecta.

# INSESGADEZ DE LOS ESTIMADORES MCO

- Bajo los supuestos S.1 a S.4 los estimadores por MCO son estimadores insesgados de los parámetros poblacionales, esto es:

$$E(\hat{\beta}_j) = \beta_j \quad \forall j = 0,1,2,\ldots,k$$

**para cualquiera de los valores del parámetro poblacional $\beta_j$**

**Insesgadez** es una descripción del estimador, esto es, en una muestra concreta, nuestro valor del estimador puede estar cerca o lejos del verdadero parámetro, pero su distribución estadística estará centrada en torno a su valor poblacional.

Pueden existir **problemas de especificación** que hagan que los estimadores por MCO no sean insesgados.

***Para $\beta_1$ en $y_i = \beta_0 + \beta_1 x + u$ :***

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(\beta_0 + \beta_1 x_i + u_i)}{\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

$$= \beta_1 + \frac{\sum_{i=1}^{n}(x_i - \bar{x})u_i}{\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

$$E(\hat{\beta}_1) \underset{\substack{\downarrow \\ S.1\ a\ S.4}}{=} \beta_1$$

Nota: En el cuadro, los valores esperados están condicionados a los valores de las variables independientes en la muestra.

# HOMOSCEDASTICIDAD VERSUS HETEROSCEDASTICIDAD

- **Supuesto 5. Homoscedasticidad (o varianza constante):**

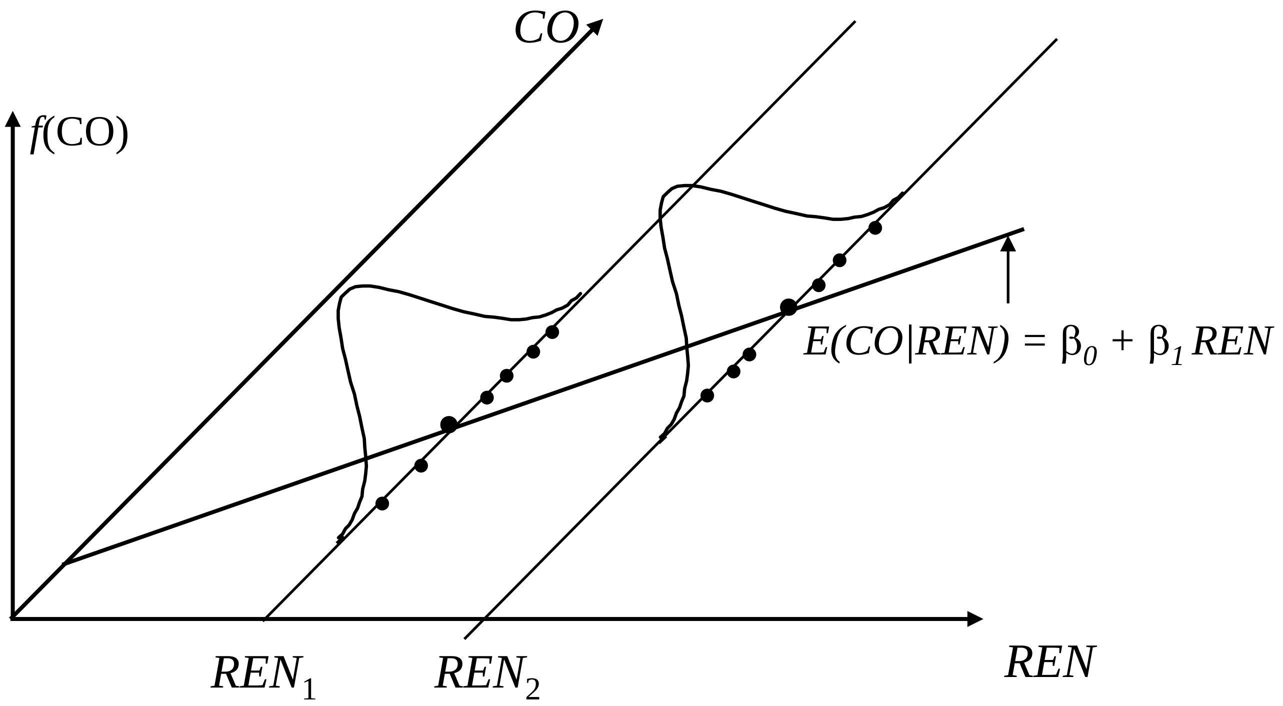
$$\mathrm{Var}(u|x_1,\ldots x_k)=\mathrm{Var}(u|\mathbf{x})=\sigma^2 \quad \text{donde } \mathbf{x}=(x_1,\ldots x_k)$$

- Este supuesto significa que la varianza del término de error *u*, condicionada a las variables explicativas del modelo es la misma para todas las combinaciones de valores de las variables explicativas. No depende de las *x's* y es constante para todas las observaciones.
- Cuando Var(*u*|**x**) depende de *x*, *no siendo constante para las distintas observaciones*, se dice que el término de error presenta **heteroscedasticidad** .
- Una forma equivalente de enunciar el supuesto de homoscedasticidad es:
- $\mathrm{Var}(y\,|x_1,\ldots x_k)=\mathrm{Var}(y\,|\mathbf{x})=\sigma^2$ **donde** $\mathbf{x}=(x_1,\ldots x_k)$

  - Por tanto,
    - ✓ *Si hay* ***homoscedasticidad ,*** *la varianza de y dado* $\mathbf{x}=x_1, x_2,\ldots x_k$ *es constante.*
    - ✓*La* ***heterocedasticidad*** *implica que* Var(*u*|**x**), *y por tanto*, Var( *y* | **x** ) ***dependen de x***.

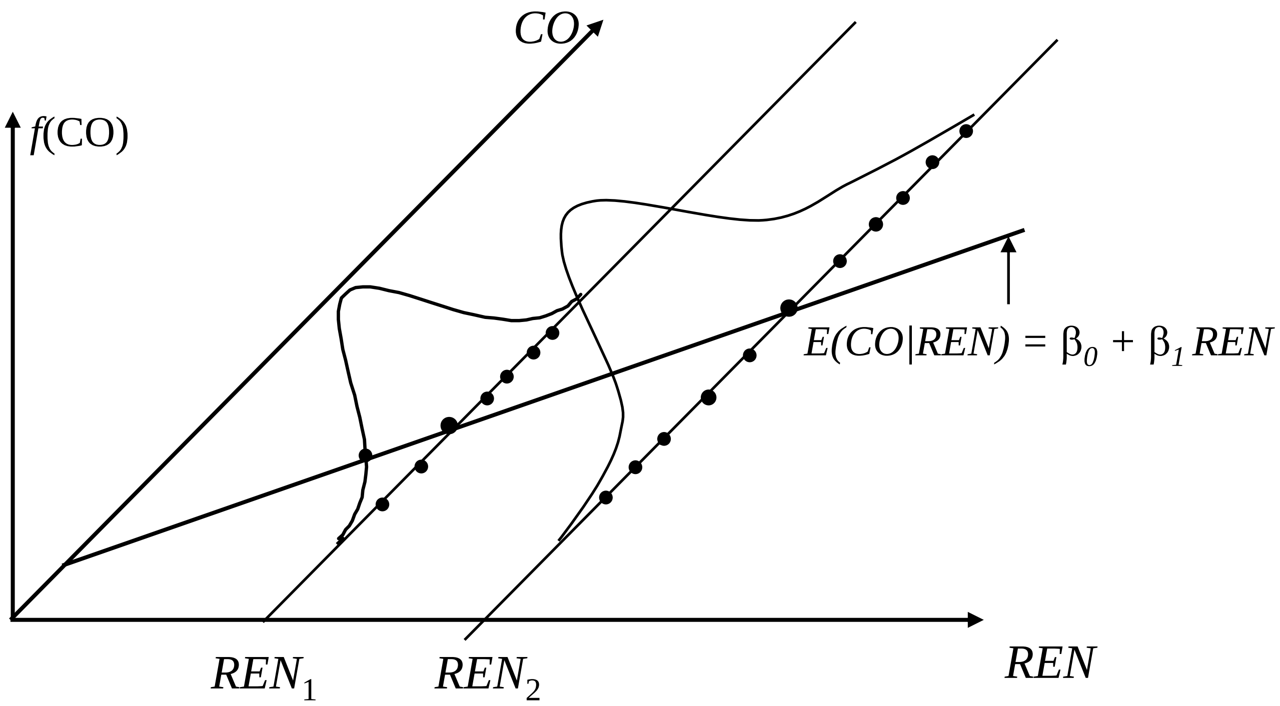
# El modelo de regresión lineal simple: la FRP

**Homoscedasticidad:** La varianza de las perturbaciones es constante para todas las observaciones.

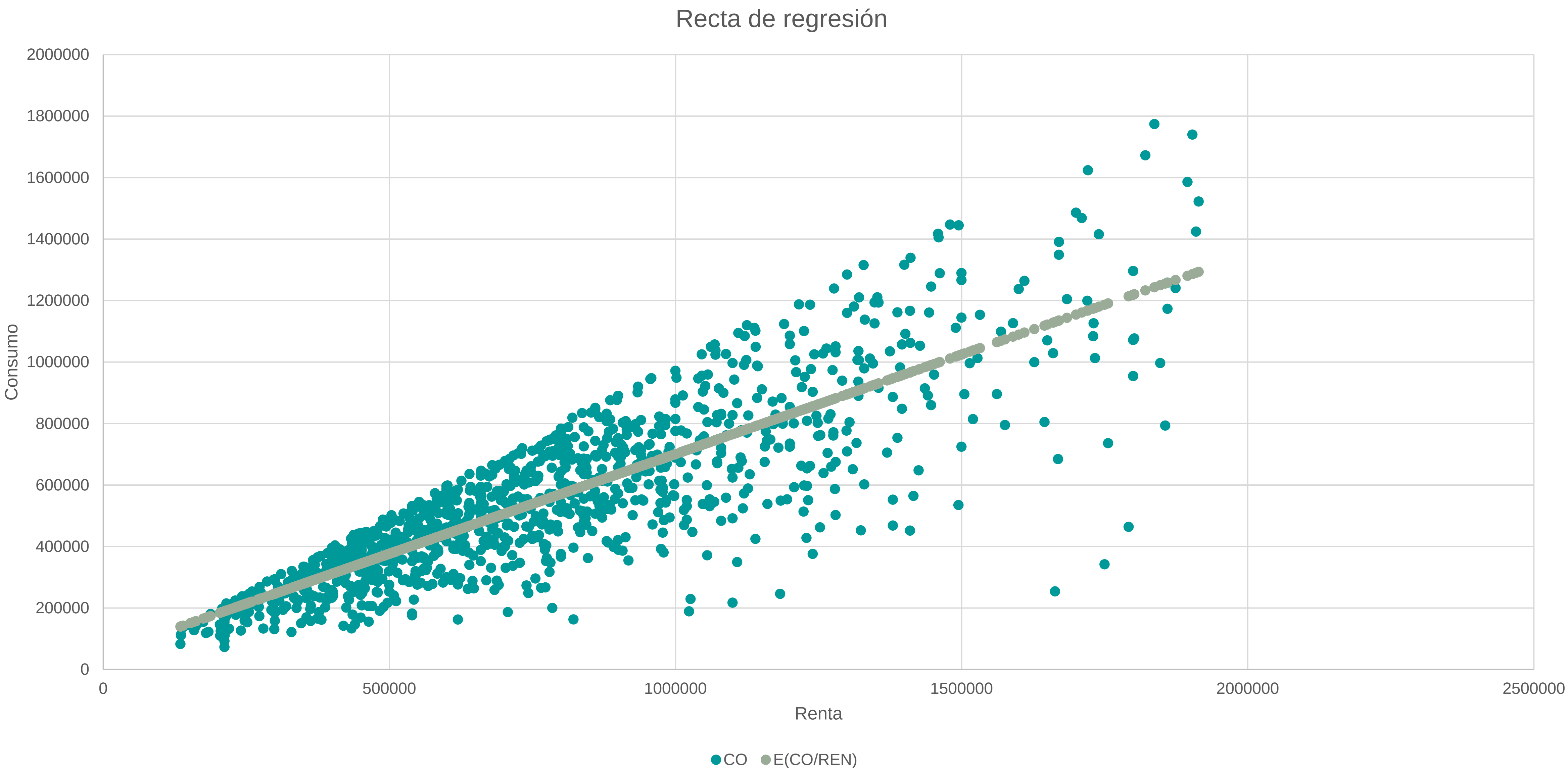


# El modelo de regresión lineal simple: la FRP

Cuando existe **heteroscedasticidad** la varianza de las perturbaciones no es constante para todas las observaciones.



# ¿Qué tenemos nosotros con nuestros datos?



Proyecto de Innovación Docente: Elaboración de materiales interactivos y multidisciplinares para favorecer el aprendizaje y evaluación en los estudios de grado y postgrado

Financiado por el Vicerrectorado de Formación Permanente, Transformación Docente y Ocupación de la Universitat de València
PROYECTO DE INNOVACION EDUCATIVA CONSOLIDADO-PIEC, UV-SFPIE_PIEC-2736595



http://www.uv.es/piclickers/

$$\begin{bmatrix} \mathbf{T} & \mathbf{doc} & \sigma^2 \\ \textit{inno} & \mathbf{I} & \mathbf{encia} \\ \mu & \textit{vación} & \mathbf{C} \end{bmatrix}$$