

LAURA DANIELA DIAZ TORRES
VIVIAN CABANZO FERNÁNDEZ
ZENNETH OLIVERO TAPIAS
CRISTIAN FELIPE MUÑOZ GUERRERO

2025-02

**BIG DATA Y MACHINE LEARNING PARA
ECONOMIA APLICADA**

Brecha salarial de género

Brecha sin controles

- Modelo:

$$\log(w) = \beta_1 + \beta_2 \text{Mujer} + u$$

- Coeficiente sexo = 0,09
→ Hombres ganan $\approx 9,4\%$ más que mujeres
- Resultado estadísticamente significativo ($p < 0,01$)
- Intervalo 95%: [0,06 – 0,11]
- La Brecha incondicional: no se ajusta por edad, educación o tipo de empleo

Brecha con controles

- Modelo ajustado:
$$\log(w) = \beta_1 + \beta_2 \text{Mujer} + \beta_3 \text{edad} + \beta_4 \text{educación} + \beta_5 \text{Tipo empleo} + \beta_6 \text{Formalidad} + \beta_7 \text{Tamaño empresa} + u$$
- Coeficiente sexo = 0,15
→ Hombres ganan $\approx 15\%$ más que mujeres
- Intervalo 95%: [0,12 – 0,16]
- La brecha persiste aún controlando factores observables, lo cual sugiere presencia de discriminación o segregación ocupacional en manera de presunción

Variables de Control (efectos esperados)

- Educación $\uparrow \rightarrow$ salarios más altos (efecto creciente)
- Formalidad $\uparrow \rightarrow +33\%$ en ingresos
- Tamaño empresa $\uparrow \rightarrow$ más grandes \rightarrow salarios más altos
- Edad efecto positivo pero no lineal

Estimación FWL

Método Frisch-Waugh-Lovell:

- Coef. res_sex = 0,1414
→ Hombres ganan $\approx 15,2\%$ más
- Altamente significativo ($t=13,07$, $p<0,001$)
- Reafirma que la brecha salarial no desaparece al controlar covariables

FWL + Bootstrap

- Intervalo 95%: [0,1197 – 0,1627]
- Confirma robustez del resultado
- Diferencia salarial significativa
→ no es azar ni supuestos estadísticos