

# Descomposicion Oaxaca-Blinder

## Explicando Brechas Salariales entre Grupos

EC3003B - Economia Laboral Aplicada

Tecnologico de Monterrey, Campus Puebla

Lunes 16 de febrero, 2025 | 3-5pm

# Contenido

- 1 Introduccion
- 2 La Descomposicion
- 3 Implementacion en Stata
- 4 Consideraciones Importantes
- 5 Aplicacion al Proyecto
- 6 Resumen

## En M01-M03 encontramos:

- Las mujeres ganan  $\approx 15\%$  menos que los hombres
- Incluso controlando por educacion y experiencia

**En M01-M03 encontramos:**

- Las mujeres ganan  $\approx 15\%$  menos que los hombres
- Incluso controlando por educacion y experiencia

**Pero... ¿que significa ese 15 %?**

- ¿Es discriminacion?
- ¿Es porque las mujeres tienen menos educacion?
- ¿Es porque trabajan en sectores que pagan menos?
- ¿Es una combinacion de factores?

En M01-M03 encontramos:

- Las mujeres ganan  $\approx 15\%$  menos que los hombres
- Incluso controlando por educacion y experiencia

Pero... ¿que significa ese 15 %?

- ¿Es discriminacion?
- ¿Es porque las mujeres tienen menos educacion?
- ¿Es porque trabajan en sectores que pagan menos?
- ¿Es una combinacion de factores?

Objetivo de hoy

Aprender a **descomponer** la brecha salarial en componentes interpretables.

Al finalizar, podras:

- ➊ Derivar la descomposicion Oaxaca-Blinder
- ➋ Interpretar los componentes “explicado” e “inexplicado”
- ➌ Implementar la descomposicion en Stata
- ➍ Discutir las limitaciones de la metodologia
- ➎ Aplicar al analisis de equidad salarial

Brecha observada:

$$\Delta = \bar{Y}_H - \bar{Y}_M$$

Brecha observada:

$$\Delta = \bar{Y}_H - \bar{Y}_M$$

Queremos separar en:

- ① **Diferencias en características** (“dotaciones”)
  - Hombres tienen mas experiencia, educacion diferente, etc.
- ② **Diferencias en retornos** (“coeficientes”)
  - Misma educacion pero “paga” diferente
  - Potencial discriminacion



Modelos separados por grupo:

$$\text{Hombres: } \bar{Y}_H = \bar{X}_H \hat{\beta}_H$$

$$\text{Mujeres: } \bar{Y}_M = \bar{X}_M \hat{\beta}_M$$

Modelos separados por grupo:

$$\text{Hombres: } \bar{Y}_H = \bar{X}_H \hat{\beta}_H$$

$$\text{Mujeres: } \bar{Y}_M = \bar{X}_M \hat{\beta}_M$$

Brecha total:

$$\Delta = \bar{Y}_H - \bar{Y}_M = \bar{X}_H \hat{\beta}_H - \bar{X}_M \hat{\beta}_M$$

Modelos separados por grupo:

$$\text{Hombres: } \bar{Y}_H = \bar{X}_H \hat{\beta}_H$$

$$\text{Mujeres: } \bar{Y}_M = \bar{X}_M \hat{\beta}_M$$

Brecha total:

$$\Delta = \bar{Y}_H - \bar{Y}_M = \bar{X}_H \hat{\beta}_H - \bar{X}_M \hat{\beta}_M$$

Truco: sumar y restar  $\bar{X}_M \hat{\beta}_H$

$$\Delta = \underbrace{(\bar{X}_H - \bar{X}_M) \hat{\beta}_H}_{\text{Explicado (Q)}} + \underbrace{\bar{X}_M (\hat{\beta}_H - \hat{\beta}_M)}_{\text{Inexplicado (U)}}$$

## Componente Explicado (Q)

$$Q = (\bar{X}_H - \bar{X}_M)\hat{\beta}_H$$

Brecha debido a **diferencias en características**.

“Si las mujeres tuvieran las mismas características que los hombres (pero sus propios retornos), cuanto cambiaria su salario?”

## Componente Explicado (Q)

$$Q = (\bar{X}_H - \bar{X}_M)\hat{\beta}_H$$

Brecha debido a **diferencias en características**.

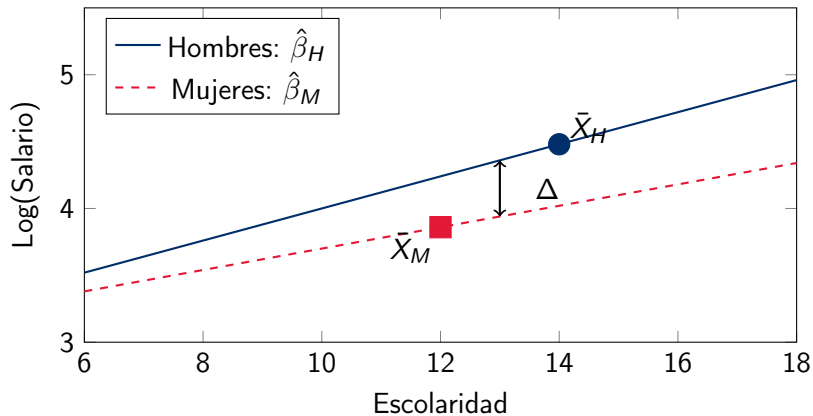
“Si las mujeres tuvieran las mismas características que los hombres (pero sus propios retornos), cuanto cambiaria su salario?”

## Componente Inexplicado (U)

$$U = \bar{X}_M(\hat{\beta}_H - \hat{\beta}_M)$$

Brecha debido a **diferencias en retornos**.

“Si las mujeres recibieran los mismos retornos que los hombres (pero sus propias características), cuanto cambiaria su salario?”



```
* Instalar si no esta
ssc install oaxaca, replace

* Descomposicion basica
oaxaca ln_salario escolaridad experiencia experiencia2, ///
      by(mujer) robust

* Descomposicion detallada (contribucion de cada variable)
oaxaca ln_salario escolaridad experiencia experiencia2, ///
      by(mujer) detail robust
```

Componente	Coeficiente	% de brecha
Brecha total	0.180	100 %
Explicado (Q)	0.050	28 %
Escolaridad	0.015	8 %
Experiencia	0.035	19 %
Inexplicado (U)	0.130	72 %

## Interpretación:

- 28 % de la brecha se explica por diferencias en educación/experiencia
- 72 % NO se explica por estas variables (potencial discriminación + omitidas)



# El Problema del Grupo de Referencia

La descomposicion depende de cual grupo usamos como referencia:

**Opcion 1: Coeficientes de hombres**

$$\Delta = (\bar{X}_H - \bar{X}_M)\hat{\beta}_H + \bar{X}_M(\hat{\beta}_H - \hat{\beta}_M)$$

**Opcion 2: Coeficientes de mujeres**

$$\Delta = (\bar{X}_H - \bar{X}_M)\hat{\beta}_M + \bar{X}_H(\hat{\beta}_H - \hat{\beta}_M)$$

# El Problema del Grupo de Referencia

La descomposicion depende de cual grupo usamos como referencia:

**Opcion 1: Coeficientes de hombres**

$$\Delta = (\bar{X}_H - \bar{X}_M)\hat{\beta}_H + \bar{X}_M(\hat{\beta}_H - \hat{\beta}_M)$$

**Opcion 2: Coeficientes de mujeres**

$$\Delta = (\bar{X}_H - \bar{X}_M)\hat{\beta}_M + \bar{X}_H(\hat{\beta}_H - \hat{\beta}_M)$$

**Solucion: Descomposicion de Neumark/Oaxaca-Ransom**

Usar coeficientes “neutrales” de un modelo pooled.

```
oaxaca ln_salario ..., by(mujer) pooled
```

## ① Inexplicado $\neq$ Discriminacion

- Puede haber variables omitidas (habilidades, preferencias)
- Sesgo de seleccion (quienes trabajan)

## ① Inexplicado $\neq$ Discriminacion

- Puede haber variables omitidas (habilidades, preferencias)
- Sesgo de seleccion (quienes trabajan)

## ② Problema de indice

- Resultados cambian con grupo de referencia
- Cambio de base en variables categoricas afecta detalle

## ① Inexplicado $\neq$ Discriminacion

- Puede haber variables omitidas (habilidades, preferencias)
- Sesgo de seleccion (quienes trabajan)

## ② Problema de indice

- Resultados cambian con grupo de referencia
- Cambio de base en variables categoricas afecta detalle

## ③ Supone modelo lineal correcto

- Si hay no linealidades, los resultados son aproximados

## ① Inexplicado $\neq$ Discriminacion

- Puede haber variables omitidas (habilidades, preferencias)
- Sesgo de seleccion (quienes trabajan)

## ② Problema de indice

- Resultados cambian con grupo de referencia
- Cambio de base en variables categoricas afecta detalle

## ③ Supone modelo lineal correcto

- Si hay no linealidades, los resultados son aproximados

## ④ Correlacion no es causalidad

- Diferencias en características pueden ser endogenas

## Usos en el proyecto de compensaciones:

- ① **Benchmark externo:** ¿Como se compara la brecha de genero de la empresa con el mercado?
- ② **Diagnostico interno:** Descomponer brechas entre:
  - Areas funcionales
  - Niveles jerarquicos
  - Antigüedad
- ③ **Identificar problemas:** Si la parte inexplicada es grande, hay que investigar

## Ejemplo: Brecha por Area

```
* Crear dummy de area (campo vs oficina)
gen campo = (area == "Operativa" | area == "Tecnica")

* Descomposicion
oaxaca ln_salario escolaridad experiencia antiguedad, ///
      by(campo) detail robust
```

### Preguntas a responder:

- ¿Personal de campo gana mas o menos que oficina?
- ¿Cuanto se explica por diferencias en educacion/experiencia?
- ¿Hay evidencia de compensacion diferencial por riesgo?



## Descomposicion O-B:

$$\Delta = \underbrace{Q}_{\text{Explicado}} + \underbrace{U}_{\text{Inexplicado}}$$

## Interpretacion:

- Q: diferencias en características
- U: diferencias en retornos

## Comandos Stata:

- `oaxaca`
- Opciones: `detail`, `pooled`

## Precaucion:

- Inexplicado  $\neq$  Discriminacion
- Variables omitidas

# ¿Preguntas?

Proxima Sesión:

## **M05: Regresión Cuantílica**

Lunes 16 de febrero, 5-7pm