

# Descomposición Oaxaca-Blinder

## Explicando Brechas Salariales entre Grupos

EC3003B - Economía Laboral Aplicada

Tecnológico de Monterrey, Campus Puebla

Lunes 16 de febrero, 2025 | 3-5pm

# Contenido

- 1 Introducción
- 2 La Descomposición
- 3 Implementación en Stata
- 4 Consideraciones Importantes
- 5 Aplicación al Proyecto
- 6 Resumen

## En M01-M03 encontramos:

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

## En M01-M03 encontramos:

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

## Pero... ¿que significa ese 15 %?

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

# Motivación: La Brecha de Género

## En M01-M03 encontramos:

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

## Pero... ¿que significa ese 15%?

- ¿Es discriminación?
- ¿Es porque las mujeres tienen menos educación?
- ¿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.

# Objetivos de Aprendizaje

Al finalizar, podrás:

- ① Derivar la descomposición Oaxaca-Blinder
- ② Interpretar los componentes “explicado” e “inexplicado”
- ③ Implementar la descomposición en Stata
- ④ Discutir las limitaciones de la metodología
- ⑤ Aplicar al análisis de equidad salarial

# Intuición de la Descomposición

**Brecha observada:**

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

# Intuición de la Descomposición

Brecha observada:

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

Queremos separar en:

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

## 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  $\bar{X}_M \hat{\beta}_H$**

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

# Interpretación de los Componentes

## 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), cuánto cambiaría su salario?”

# Interpretación de los Componentes

## 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), cuánto cambiaría su salario?”

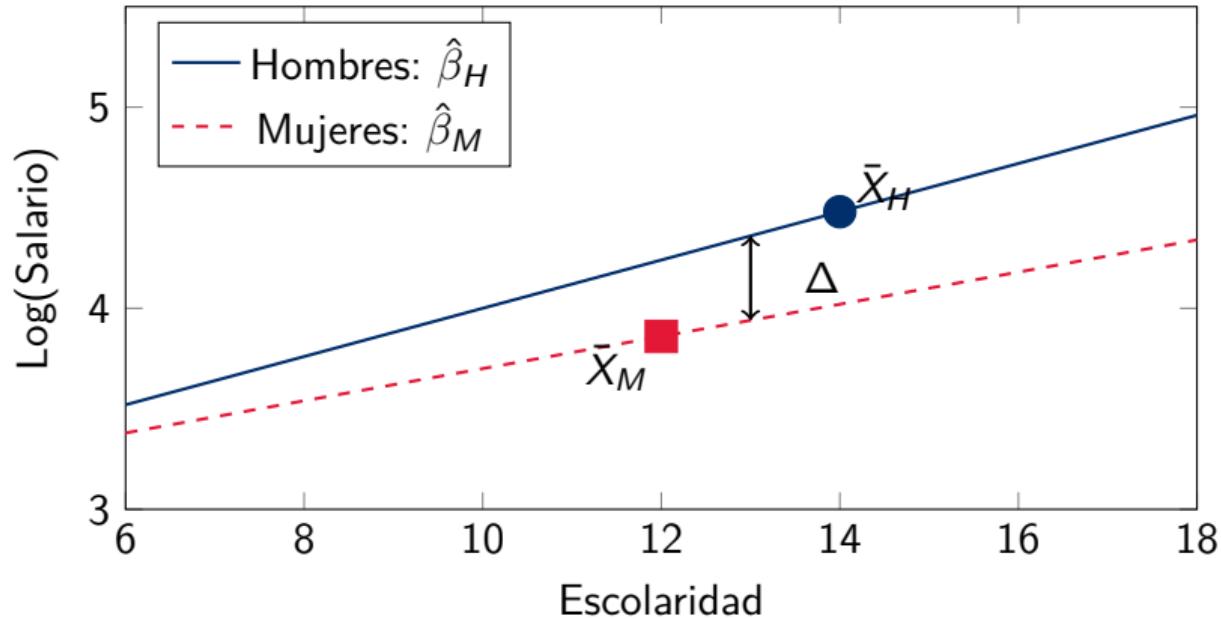
## Componente Inexplicado (U)

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

Brecha debido a **diferencias en retornos**.

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

# Visualización



# Comando oaxaca en Stata

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

* Descomposición básica
oaxaca ln_salario escolaridad experiencia experiencia2, ///
    by(mujer) robust

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

# Resultados Típicos

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 descomposición 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 descomposición 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

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

Solución: Descomposición de Neumark/Oaxaca-Ransom

Usar coeficientes “neutrales” de un modelo pooled.

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

## ① Inexplicado ≠ Discriminación

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

# Limitaciones de la Metodología

## ① Inexplicado ≠ Discriminación

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

## ② Problema de índice

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

# Limitaciones de la Metodología

## ① Inexplicado ≠ Discriminación

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

## ② Problema de índice

- 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

# Limitaciones de la Metodología

## ① Inexplicado ≠ Discriminación

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

## ② Problema de índice

- 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 Correlación no es causalidad

- Diferencias en características pueden ser endogenas

## Usos en el proyecto de compensaciones:

- ① **Benchmark externo:** ¿Cómo se compara la brecha de género de la empresa con el mercado?
- ② **Diagnóstico interno:** Descomponer brechas entre:
  - Áreas funcionales
  - Niveles jerárquicos
  - Antiguedad
- ③ **Identificar problemas:** Si la parte inexplicada es grande, hay que investigar

## Ejemplo: Brecha por Área

```
* Crear dummy de área (campo vs oficina)
gen campo = (área == "Operativa" | área == "Técnica")

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

### Preguntas a responder:

- ¿Personal de campo gana más o menos que oficina?
- ¿Cuánto se explica por diferencias en educación/experiencia?
- ¿Hay evidencia de compensación diferencial por riesgo?

# Resumen de la Sesión

## Descomposición O-B:

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

## Interpretación:

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

## Comandos Stata:

- oaxaca
- Opciones: detail, pooled

## Precaucion:

- Inexplicado  $\neq$  Discriminación
- Variables omitidas

# ¿Preguntas?

Próxima Sesión:

**M05: Regresión Cuantilica**

Lunes 16 de febrero, 5-7pm