

Problem Set 1: Predicting Income

Grupo 3

Gustavo Adolfo Castillo Álvarez (201812166), Alexander Almeida Ramírez (),
Jorge Luis Congacha Yunda (201920042) y
Jaime Orlando Buitrago González (200612390)

03 de marzo de 2024

Introducción

Datos

De las 32177 observaciones, se puede observar la tasa de valores perdidos de las variables de interés:

Perfiles de salario por edad

Brechas de ingreso por sexo

Predicción de salarios

La meta es predecir la brecha salarial del logaritmo del ingreso. Comenzamos estimando el modelo más sencillo de todos, es decir, el modelo univariado:

$$\ln(w) = \beta_1 + \beta_2 \text{Mujer} + u \quad (1)$$

El coeficiente $\beta_2 < 0$ indica que las mujeres, en promedio y ceteris paribus, reciben un salario mensual 0.8657713 menos ingresos que los hombres.

Estimación con FWL

En la primera etapa, *partialling-out*, ejecutamos dos regresiones. Suponiendo que la variable de interés es la dicotómica de *Mujer*, entonces en la primera regresión buscamos estimar `woman~x1+x2+...`, donde las `xi` son todas aquellas demás variables de control para corregir el potencial sesgo de variable omitida. Luego nos quedamos con los residuales `woman_res`, y ejecutamos una segunda regresión en la que estimemos `log_wage~x1+x2+...`, y guardamos estos residuales, `log_wage_res`.

Finalmente ejecutamos la segunda regresión univariada `log_wage_res~woman_res` y en principio deberíamos obtener el mismo coeficiente de haber ejecutado el modelo completo.

Table 1: Estimando brecha de género

	<i>Dependent variable:</i>	
	Ln Salario	
	(1)	(2)
age		0.012*** (0.001)
womanMujer	-0.144*** (0.016)	-0.124*** (0.014)
relab		0.126*** (0.020)
Constant	14.076*** (0.011)	13.898*** (0.192)
Observations	9,964	9,963
R ²	0.009	0.453
Adjusted R ²	0.008	0.448
Residual Std. Error	0.775 (df = 9962)	0.579 (df = 9876)
F Statistic	86.083*** (df = 1; 9962)	94.912*** (df = 86; 9876)

Note:

*p<0.1; ** p<0.05; *** p<0.01
 Controles: oficio, maxEduclevel

Referencias bibliográficas

Ejemplos

Para incrustar código y resultados de la consola

```
summary(cars)
```

```
##      speed      dist
##  Min.   : 4.0    Min.   :  2.00
##  1st Qu.:12.0    1st Qu.: 26.00
##  Median :15.0    Median : 36.00
##  Mean   :15.4    Mean   : 42.98
##  3rd Qu.:19.0    3rd Qu.: 56.00
##  Max.   :25.0    Max.   :120.00
```

Para incluir ecuaciones

$$w = f(X) + u$$

Para incluir gráficas

