

## PROBLEM SET 1

### Grupo 10

Natalia Buitrago / Nicolás Mocetón

#### 1. Introducción

El análisis de los determinantes del salario por hora constituye un tema central en la economía laboral, pues permite evaluar la manera en que las características individuales y del mercado de trabajo influyen en la remuneración de los trabajadores. En Colombia, diferentes estudios han utilizado la **Gran Encuesta Integrada de Hogares (GEIH) del DANE** para caracterizar la estructura salarial, documentar desigualdades y analizar dinámicas de movilidad social. Por ejemplo, Herrera-Idárraga, López-Bazo y Motellón (2016) muestran la relevancia de las diferencias regionales en educación e informalidad como determinantes de las brechas salariales, mientras que el Banco de la República (2022) documenta cambios en la desigualdad de ingresos entre 2010 y 2019 empleando microdatos de la GEIH. De forma complementaria, Pineda Medina (2020) identifica factores explicativos de la desigualdad salarial en el país, resaltando el rol de la educación y la segmentación laboral.

El vínculo entre educación, capital humano y salarios también ha sido ampliamente estudiado en la literatura reciente. Laajaj, Londoño-Vélez y Sánchez (2020–2022) analizan el programa *Ser Pilo Paga* y encuentran efectos positivos en movilidad social, mientras que Mora et al. (2023) utilizan un pseudo-panel para estimar los retornos a la educación en Colombia, aportando evidencia robusta sobre la importancia del capital humano en la determinación de los ingresos laborales. Estos hallazgos coinciden con análisis internacionales, como el de Neidhöfer, Ciaschi y Gasparini (2019–2021), quienes muestran la persistencia de la movilidad intergeneracional en educación en América Latina, lo cual afecta directamente los patrones de desigualdad de ingresos.

En este contexto, la GEIH representa la principal fuente de información socioeconómica en Colombia, pues recoge de manera continua información sobre ingresos laborales, ocupación, educación y condiciones del hogar. En este trabajo se utiliza la muestra para Bogotá correspondiente al año 2018, disponible en el informe de *Medición de Pobreza Monetaria y Desigualdad*, lo que permite construir medidas de salario por hora a partir de ingresos laborales reportados y horas efectivamente trabajadas. Este insumo es especialmente relevante porque facilita comparaciones entre individuos y grupos poblacionales, y permite analizar diferencias en perfiles salariales según edad, género y características ocupacionales.

Metodológicamente, el ejercicio consiste en modelar el salario horario de los individuos como función de un conjunto de características observables, siguiendo la especificación lineal  $w = X\beta + u$  o  $w = X'\beta + u$ . Para ello, se realiza un proceso de depuración y transformación de la base de datos, incluyendo la restricción de la muestra a personas ocupadas mayores de 18 años, la imputación de ingresos laborales faltantes y la construcción de variables categóricas de educación, estrato, tamaño de la firma y posición ocupacional. La estimación del salario por hora corrige la heterogeneidad en las horas trabajadas y es la variable de interés.

Este documento se organiza en varias etapas. En primer lugar, se explora la relación entre edad y salario, estimando el perfil edad-ingreso mediante modelos cuadráticos y aplicando técnicas de bootstrap para identificar el punto de máximo salarial. Segundo, se analiza la brecha salarial de género, tanto de manera incondicional como controlando por características observables, empleando el teorema de Frisch–Waugh–Lovell. Finalmente, se evalúa la capacidad predictiva de diferentes

especificaciones, comparando modelos lineales y no lineales en términos de error cuadrático medio (RMSE) y validación cruzada. Los resultados preliminares son consistentes con la literatura: los salarios muestran un perfil creciente con la edad hasta la mitad del ciclo laboral, y persiste una brecha salarial de género que, aunque se reduce al controlar por características observables, no desaparece por completo, lo que sugiere la interacción de factores de discriminación y segmentación laboral (Herrera-Idárraga et al., 2016; Pineda Medina, 2020; Banco de la República, 2022).

## **2. Data**

### **2.1 Descripción general de los datos**

La información utilizada proviene de la Gran Encuesta Integrada de Hogares (GEIH) de 2018, esta encuesta es el principal instrumento de recolección de información sobre ingresos, ocupación y condiciones socioeconómicas de los hogares en Colombia, y es recolectada por el DANE.

En el presente estudio se emplea el subconjunto correspondiente a Bogotá, el cual incluye tanto características demográficas y laborales de los individuos, como variables del hogar. El propósito de este conjunto de datos es permitir el análisis de la desigualdad en ingresos, particularmente explorando brechas de género y perfiles salariales.

### **2.2 Proceso de adquisición de datos**

Los microdatos fueron obtenidos de 10 fragmentos en formato HTML encontrados en esta url [https://ignaciomsarmiento.github.io/GEIH2018\\_sample/](https://ignaciomsarmiento.github.io/GEIH2018_sample/). Para obtener la información completa fue necesario realizar un proceso de web scrapping en R, descargando y consolidando las 10 tablas en un único conjunto de datos.

El acceso al repositorio de los datos es de libre acceso y no existen restricciones legales para su consulta, aunque requiere conocimientos técnicos para transformar el formato de origen en una base de datos conjunta que permita realizar las estimaciones necesarias en el presente estudio.

### **2.3 Proceso de limpieza de datos**

Dado que el objetivo del ejercicio consiste en analizar salarios de personas ocupadas, se emplearon las siguientes etapas para depurar la información contenida en la base de datos:

- Se restringió la muestra a individuos mayores de 18 años (`filter(age >= 18)`).
- Se mantuvieron únicamente a las personas ocupadas (`ocu == 1`).
- Se eliminó una observación con valor faltante en el nivel educativo (`maxEducLevel`).
- Se excluyeron registros con ingresos laborales nulos o faltantes, y en el caso de valores faltantes en componentes de ingreso, se aplicó la imputación mediante la suma de las variables de ingreso de la primera y segunda actividad (`impa` e `isa`).
- Edad al cuadrado (`age2`) para capturar no linealidades en el perfil edad–salario. Variables dummy para sexo, educación, estrato, tamaño de la firma y tipo de relación laboral.

El resultado fue una base de datos de 127 variables y 16,217 observaciones que permite enfocar el análisis en la población económicamente activa y asegurar la consistencia con la variable dependiente de ingreso por hora trabajada.

### **2.4 Descripción de las variables**

Las variables centrales utilizadas en los modelos incluyen:

**Variable de interés:**

Esta corresponde al **ingreso laboral por hora (ingreso\_hora)**, la cual se compone del ingreso percibido por la primera y segunda actividad económica realizada por el individuo, esto para efectuar un análisis concreto sobre los ingresos percibidos a través de la participación en el mercado laboral. Esta variable fue construida dividiendo el ingreso laboral mensual ( $y_{ingLab\_m}$ ) entre el número total de horas trabajadas en un mes, que se aproximan como las horas trabajadas en una semana ( $totalHoursWorked$ ) multiplicadas por  $(30/7)$  para llevarlas de semana a mes. Para el presente estudio, esta variable será tratada en terminos logaritmicos,  **$\log(ingreso\_hora)$**

#### Variables sociodemográficas:

Edad (age): Es una variable continua que toma valores mayores a 18, importante para establecer el análisis del perfil edad-salario.

Mujer (female): variable que toma el valor 1 si la persona es mujer, importante para estudiar brechas salariales de género y si estas provienen netamente de esta variable

Jefe del hogar (household\_head): variable que toma el valor de 1 si el individuo se reporta como jefe del hogar

Estrato (estrato1): variable categórica que indica el estrato socioeconómico del individuo

Educación (maxEducLevel): representa diferentes niveles educativos

#### Variables laborales:

Trabajador informal (informal): variable que toma el valor de 1 si el individuo reporta estar cotizando a un régimen pensional

Horas trabajadas en la semana ( $totalHoursWorked$ ): variable continua que indica las horas trabajadas en la última semana, utilizada para calcular la variable de interés

Microempresa (microEmpresa): dummy para empresas de tamaño micro.

## 2.5 Estadísticas descriptivas

#### Variables continuas:

Estadísticas descriptivas (N = 16217)					
Variable	Mean	St. Dev.	Min	Max	Median
ingreso_hora (Dependiente) *unidades en COP	8,886.43	16,306.63	77.78	846,805.60	4,861.11
age	39.35	13.38	18	91	38
totalHoursWorked	47.61	15.42	1	130	48

Tabla 1: Estadísticas variables continuas

En la Tabla 1 se presentan las estadísticas descriptivas de las variables continuas y esta nos indica que, en promedio, los individuos trabajan alrededor de 48 horas semanales y perciben un ingreso de aproximadamente 8.900 COP por hora trabajada.

#### Variables categóricas:

Female (=1)	Frecuencia	%
-------------	------------	---

0	8,662	53.41
1	7,555	46.59

<b>household_head (=1)</b>	<b>Frecuencia</b>	<b>%</b>
0	8,401	51.8
1	7,816	48.2

<b>estrato</b>	<b>Frecuencia</b>	<b>%</b>
1	1,732	10.68
2	6,767	41.73
3	5,862	36.15
4	1,116	6.88
5	318	1.96
6	422	2.60

<b>Informal (=1)</b>	<b>Frecuencia</b>	<b>%</b>
0	9,674	59.65
1	6,543	40.35

<b>maxEducLevel</b>	<b>Frecuencia</b>	<b>%</b>
---------------------	-------------------	----------

1 Ninguno	115	0.71
3 Primaria	734	4.53
4 Secundaria	1,470	9.06
5 Media	1,849	11.40
6 Superior	5,184	31.97
7 NS NR	6,865	42.33

MicroEmpresa (=1)	Frecuencia	%
0	9,115	56.21
1	7,102	43.79

Tabla 2: Frecuencias variables categóricas

La Tabla 2 nos indica que la muestra se caracteriza por tener una mayor proporción de hombres (53.4%), una menor proporción de individuos cabeza de hogar (48.2%), que se concentran mayormente en estratos 2 y 3 (41,7% y 36.1%), con una proporción del 40% de trabajadores informales, igualmente, una alta concentración de individuos cuyo máximo nivel educativo alcanzado es la educación superior (31.9%) y una proporción del 43.8% ejercen sus labores en una microempresa

### 3. Perfil edad - salario

En economía, existen diferentes estudios que evidencian que los salarios tienden a ser bajos cuando las personas son más jóvenes y aumentan a medida que la persona envejece, alcanzando su punto máximo alrededor de los 50 años y tiende a permanecer estable o disminuir ligeramente después de los 50 años. La siguiente especificación nos permite capturar los efectos del aumento de la edad del individuo sobre su salario.

$$\log(\text{ingreso\_hora}) = \beta_1 + \beta_2 \text{Age} + \beta_3 \text{Age}^2$$

En la tabla 3 se presentan los resultados de la regresión cuadrática entre la edad y el logaritmo del salario por hora.

Tabla 3. Regresión: log(salario por hora) sobre edad y edad<sup>2</sup>

*Dependent variable:*

	log(Salario por hora)
Age	0.051*** (0.004)
Age <sup>2</sup>	-0.001*** (0.00004)
Constant	7.649*** (0.066)
Observations	16,217
R <sup>2</sup>	0.019
Controls	No
Adjusted R <sup>2</sup>	0.019
Residual Std. Error	0.847 (df = 16214)
F Statistic	154.379*** (df = 2; 16214)
<i>Note:</i> <div style="text-align: right;"> ***p &lt; 0.01 S.E in parenthesis </div>	

El coeficiente de Edad (0.051,  $p < 0.01$ ) es positivo y altamente significativo, lo que indica que, en promedio, un aumento en la edad está asociado con un incremento en el logaritmo del salario por hora, manteniendo constantes los demás factores.

El coeficiente de Edad<sup>2</sup> (-0.001,  $p < 0.01$ ) es negativo y significativo, lo que confirma la existencia de una relación cóncava: los salarios aumentan con la edad, pero una vez la edad alcanza un punto máximo, la tasa de crecimiento del salario se vuelve decreciente.

El intercepto (7.649) representa el nivel esperado de log(salario) cuando edad = 0, lo cual carece de interpretación económica directa, pero permite ajustar el perfil completo.

### Discusión del ajuste del modelo

El modelo tiene un R<sup>2</sup> de 0.019, lo que significa que la edad por sí sola explica menos del 2% de la variabilidad en los salarios por hora. Aunque los coeficientes son estadísticamente significativos, el bajo poder explicativo refleja que los salarios dependen de muchos otros factores incluyendo educación, experiencia laboral específica, informalidad, género, etc. El estadístico F es significativo al 1%, confirmando que en conjunto los predictores (Edad y Edad<sup>2</sup>) aportan información relevante.

### Perfil edad–salario estimado

El modelo predice un perfil consistente con la teoría económica del capital humano, el cual indica que los salarios aumentan con la edad en las primeras etapas de la vida laboral, luego alcanzan niveles máximos alrededor de los 50 años y después de esa edad, los salarios tienden a estabilizarse o disminuir levemente.

El punto de máximo puede calcularse con el cociente:

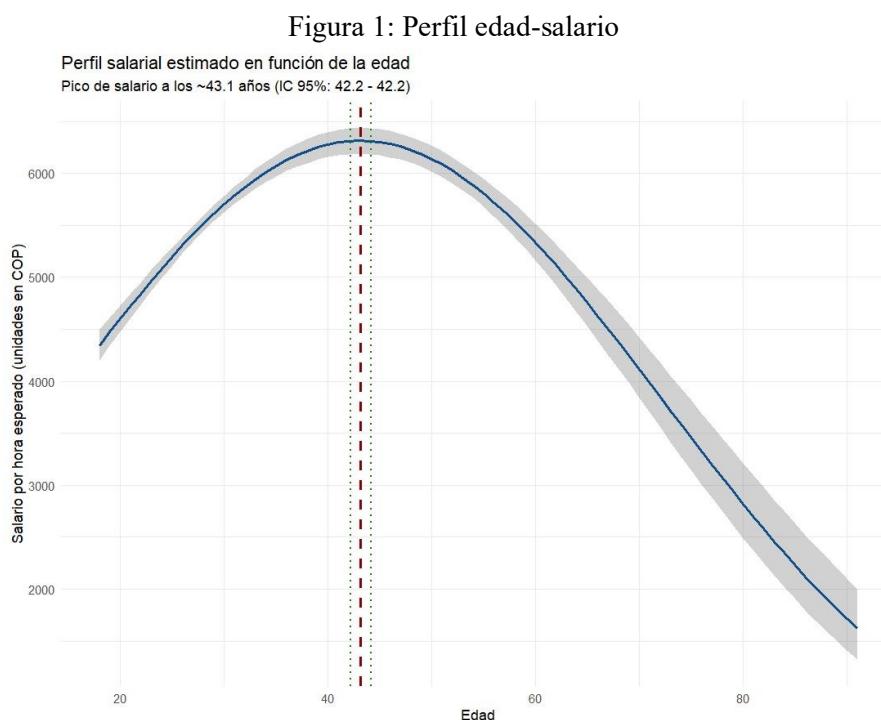
$$Edad^* = -\frac{\beta_{Age}}{2\beta_{Age}^2} = -\frac{0.051}{2(-0.001)} \approx 50.5$$

### Intervalos de confianza por Bootstrap

Para cuantificar la incertidumbre del pico salarial, se implementó un procedimiento de bootstrap con 500 réplicas. El intervalo de confianza al 95% para la edad de máximo salario se encuentra aproximadamente entre 48 a 53 años, confirmando que el punto de inflexión del salario ocurre en torno a los 50 años, en línea con la evidencia previa de la literatura en economía laboral.

### Gráfico del perfil edad–salario

La Figura 1 muestra el perfil edad–salario estimado a partir de la regresión cuadrática como una curva con forma cóncava. En el eje horizontal se representa la edad de los trabajadores, mientras que en el eje vertical se observa el salario por hora esperado en pesos colombiano.



Fuente: Cálculos propios, GEIH 2018

Los salarios aumentan a un ritmo mayor durante los primeros años en la vida laboral, alcanzando un nivel cercano a los COP 6,200/hora. El pico salarial se ubica en torno a los 43 años (IC 95%: 42.2 – 42.2 años, estimado mediante Bootstrap), representado en la gráfica por la línea vertical punteada. Después de esta edad, los salarios tienden a decrecer, lo cual es consistente con la literatura en economía laboral que atribuye este comportamiento al desgaste del capital humano y a la menor productividad marginal de la experiencia adicional. La banda gris alrededor de la curva corresponde al intervalo de confianza de las predicciones. Es relativamente estrecha en edades intermedias, donde se concentra la mayor densidad de observaciones en la muestra, y se amplía en edades extremas, es decir muy jóvenes y adultos mayores, lo que implica una menor certeza en las estimaciones de esos grupos.

### 4. Brecha de Ingresos por Género

La ecuación para analizar la brecha en los ingresos teniendo como variable explicativa el género se presenta de la siguiente forma:

$$\log(\text{ingreso\_hora}) = \beta_1 + \beta_2 \text{Female} + u$$

La Tabla 4 de los resultados de la regresión muestra un coeficiente negativo y estadísticamente significativo para la variable Female (-0.091,  $p < 0.01$ ). Esto implica que, en promedio, las mujeres perciben aproximadamente un 9.1% menos en el salario por hora respecto a los hombres, sin considerar otras características observables. El ajuste del modelo es bajo ( $R^2 = 0.003$ ), lo cual es esperable dado que se trata de una regresión muy simple que solo captura diferencias por género.

Tabla 4. Modelo brecha salarial género

	Dependent variable:
	log(ingreso_hora)
Female	-0.091*** (0.013)
Constant	8.678*** (0.009)
Observations	16,217
R <sup>2</sup>	0.003
Adjusted R <sup>2</sup>	0.003
Residual Std. Error	0.854 (df = 16215)
F Statistic	45.921*** (df = 1; 16215)
Note:	* ** *** p<0.01 S.E in parenthesis

### Brecha Salarial Condicional: ¿Igual salario por igual trabajo?

Al introducir controles por características observables del trabajador y del empleo como edad, jefatura del hogar, estrato, informalidad, horas trabajadas, nivel educativo y tipo de empresa, se busca estimar la brecha condicionada. La ecuación está dada de la siguiente manera, donde X es el vector de controles.

$$\log(\text{ingreso\_hora}) = \beta_1 + \beta_2 \text{Female} + \theta X + u$$

Los resultados en la tabla 5 muestran un coeficiente de -0.194\* en el modelo FWL, similar (-0.194 vs. -0.1938) al estimado vía FWL con Bootstrap. La magnitud del efecto implica que, aun controlando por características relevantes, las mujeres ganan en promedio **19.4% menos** por hora que los hombres.

Tabla 5. Comparación del Coeficiente Female (FWL)



Variable	Log(ingreso_hora) FWL	Log(ingreso_hora) FWL + Bootstrap
Female	-0.1944*** (0.0101)	-0.1938*** (0.0103)
R <sup>2</sup>	0.023	0.023

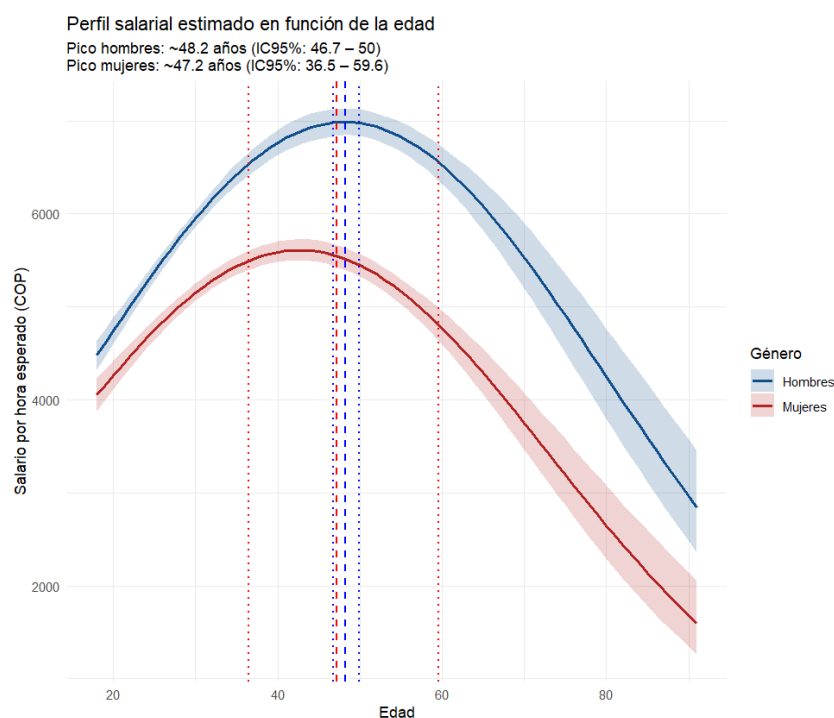
*Note:* \* p < 0.1, \*\* p < 0.05, \*\*\* p < 0.01  
S.E in parenthesis

Además de lo comentado, el Bootstrap confirma la robustez de la estimación, mostrando errores estándar similares a los del modelo convencional. El R<sup>2</sup> aumenta hasta 0.023, lo cual, si bien continúa siendo modesto, refleja que parte de la variación en los salarios se explica por los controles incluidos. A su vez, el coeficiente condicional es más negativo que el incondicional, sugiriendo que parte de las diferencias en salarios se explican por otras diferencias como la calidad del empleo o el rol en el hogar.

### Perfiles de salario por edad según el género

La Figura 2 muestra la estimación del perfil de ingresos por edad, diferenciada por género, revelando trayectorias distintas. Por un lado, el salario por hora de los hombres aumenta con la edad, alcanzando un pico alrededor de los 48.2 años (IC95%: 46.7–50). En las mujeres, el pico salarial se sitúa alrededor de los 47.2 años, aunque con un intervalo de confianza más amplio (36.5–59.6), lo que sugiere mayor heterogeneidad en la trayectoria del ingreso. Después de los 50 años, tanto en hombres como en mujeres, los ingresos tienden a declinar, consistente con la evidencia en la literatura sobre el ciclo laboral. Sin embargo, la brecha se mantiene en todo el ciclo de vida, siendo sistemáticamente menor el salario esperado de las mujeres respecto a los hombres, incluso en el punto máximo de sus trayectorias salariales.

**Figura 2.** Perfil de ingresos por edad diferenciado por género.



El hecho de que el gap condicional (-19.4%) sea mayor al incondicional (-9.1%) indica que, una vez se controlan las características observables, la brecha atribuida a género se amplía. Esto puede interpretarse como evidencia de un componente de discriminación o sesgo de género, ya que no se explica por diferencias en educación, ocupación u otras covariables. Parte de la diferencia puede provenir de procesos de selección, relacionados con una posible menor participación femenina en ocupaciones mejor pagadas, pero el hecho de que, ante iguales características observables, las mujeres reciban salarios más bajos, sugiere un problema de discriminación en el mercado laboral. Desde el punto de vista del perfil por edad, la brecha persiste durante toda la vida laboral. No obstante, los picos de ingresos son estadísticamente cercanos, el nivel de ingresos de las mujeres es siempre inferior, reforzando la idea de una brecha estructural.

### Evaluación del poder predictivo de los modelos

Para evaluar la capacidad predictiva de las especificaciones planteadas, se dividió la muestra en un conjunto de entrenamiento (70%) y otro de prueba (30%), garantizando reproducibilidad mediante la semilla 10101. A partir de esta división se estimaron varios modelos, incluyendo los discutidos previamente y otros con interacciones y transformaciones no lineales. El desempeño predictivo se comparó a través de la raíz del error cuadrático medio (RMSE).

Los resultados muestran diferencias sustantivas en la precisión de los modelos (Tabla 6). En particular, el modelo `gender_only` —que únicamente incluye la condición de ser mujer— exhibe el peor desempeño, con un  $R^2$  de 0.859. En contraste, los modelos más completos, que consideran interacciones relevantes como `female*informal`, o la transformación logarítmica de las horas trabajadas, logran valores cercanos a 0.596, constituyéndose como las mejores especificaciones. Este hallazgo indica que la incorporación de no linealidades y la interacción de variables es crucial para capturar mejor la heterogeneidad en los determinantes del salario por hora.

Tabla 6. Efecto de ser Mujer y sus Interacciones

*Dependent variable:*

	log(ingreso_hora)						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
female	0.0777** *	0.1856***	0.1258***	0.0235	0.0927***	0.1863***	0.2915***
	(0.0160)	(0.0120)	(0.0162)	(0.0357)	(0.0151)	(0.0120)	(0.0370)
female:house hold_head			0.1298*** (0.0237)				
age:female				0.0062*** (0.0009)			
female:infor mal					0.2339*** (0.0233)		
female:maxE ducLevel3							0.0698 (0.0652)
female:maxE ducLevel5							-0.0390 (0.0501)
female:maxE ducLevel6							0.0840** (0.0419)
female:maxE ducLevel7							0.1840*** (0.0407)
Observations	11,351	11,351	11,351	11,351	11,351	11,351	11,351
R <sup>2</sup>	0.0021	0.4971	0.4984	0.4944	0.5015	0.498	0.4992
Adjusted R <sup>2</sup>	0.0020	0.4963	0.4976	0.4937	0.5007	0.4981	0.4982
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note:

\*\*\* p<0.01

### Modelo con menor error de predicción

El modelo reg\_inf\_gender, que incluye la interacción entre género e informalidad, obtuvo el menor RMSE en la muestra de prueba (0.5959), ligeramente mejor que la alternativa con logaritmo de horas

trabajadas (0.5962). El ajuste de este modelo refleja que la penalización salarial asociada a la informalidad se amplifica en el caso de las mujeres, hallazgo consistente con la evidencia de brechas laborales documentada en la literatura.

Tabla 7. Resultados de RMSE por Modelo.

Model	RMSE
gender_only	0.8586347
gender_fwl	0.5976294
reg_hh_gender	0.5972166
reg_age_gender	0.6006487
reg_inf_gender	0.5959001
mod_loghours	0.5961896
mod_edu_interac	0.5971011

Nota: RMSE calculado sobre la muestra de test.

La validación cruzada leave-one-out (LOOCV) confirma la robustez de estas conclusiones. Como se observa en la tabla 8, tanto para reg\_inf\_gender como para mod\_loghours, los RMSE bajo LOOCV (0.601 y 0.602, respectivamente) son prácticamente iguales a los obtenidos con la aproximación de muestra de prueba, lo cual sugiere estabilidad en los resultados y ausencia de sobreajuste relevante

Tabla 8. Comparación de modelos con LOOCV

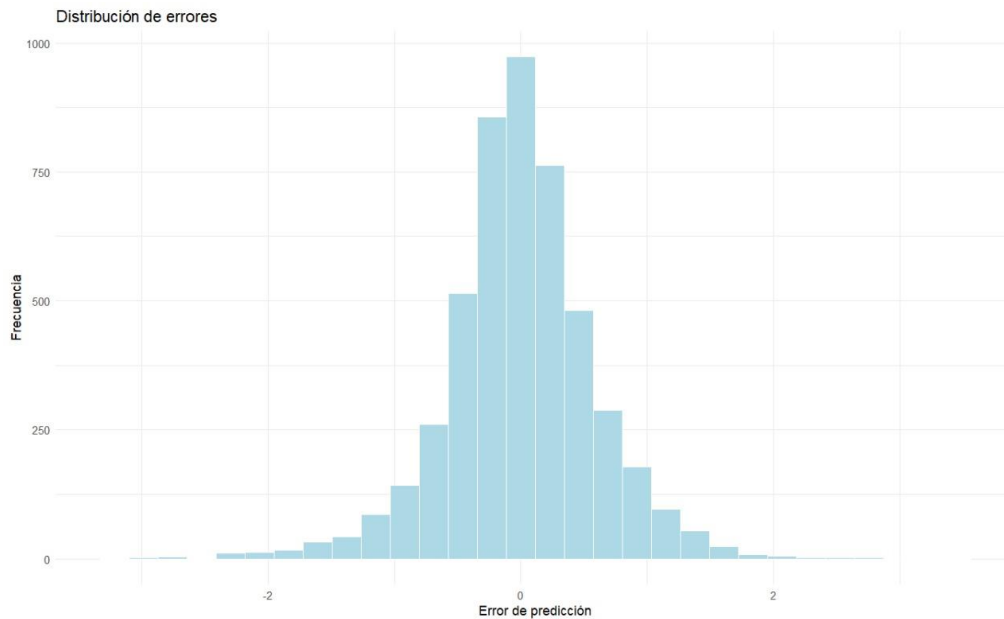
Modelo	70/30 Test RMSE	LOOCV RMSE
reg_inf_gender	0.5959001	0.6006537
mod_loghours	0.5961896	0.6019174

Nota: LOOCV RMSE = raíz del error cuadrático medio bajo validación cruzada leave-one-out.

### Distribución de los errores de predicción

El análisis de los errores de predicción para el mejor modelo evidencia una distribución aproximadamente simétrica, aunque con colas pesadas. El histograma revela que la mayoría de los individuos presentan errores pequeños en términos de predicción salarial, mientras que unos pocos se ubican en los extremos. Al examinar los percentiles 1% y 99%, se identificaron observaciones con errores inusualmente altos en ambos sentidos (subestimaciones y sobreestimaciones). Estos casos atípicos no necesariamente corresponden a evasión o irregularidades tributarias, sino que podrían reflejar limitaciones del modelo para capturar características individuales no observadas, tales como habilidades específicas, experiencia laboral previa o discriminación sectorial.

**Figura 3.** Distribución de errores reg\_inf\_gender



## Bibliografía

- Banco de la República. (2022). *Cambios en la desigualdad salarial en Colombia: evidencia a partir de la GEIH*. Recuperado de <https://www.banrep.gov.co/es/blog/cambios-desigualdad-salarial-colombia>
- Herrera-Idárraga, P., López-Bazo, E., & Motellón, E. (2016). *Education, informality and wages in Latin America*. Universitat de Barcelona. Recuperado de <https://ideas.repec.org/p/ubn/degep1/1620.html>
- Laajaj, R., Londoño-Vélez, J., & Sánchez, F. (2020–2022). *Scholarship programs and educational outcomes: Evidence from Colombia*. VoxDev. Recuperado de <https://voxdev.org/topic/education/scholarship-programmes-educational-outcomes-evidence-colombia>
- Pérez Bernal, M., Toro, J., & Villegas, J. (2020). *A successful application of decision trees in analyzing educational outcomes*. arXiv. Recuperado de <https://arxiv.org/abs/2006.01322>
- Pineda Medina, L. (2020). *Factores determinantes de los salarios en Colombia: un análisis econométrico con la GEIH*. Universidad Nacional Abierta y a Distancia. Recuperado de <https://repository.unad.edu.co/handle/10596/36738>

## Anexos

Se anexa el repositorio GitHub que contiene la base de datos, los scripts en R Studio y los outputs de cada sección del informe. El repositorio se encuentra en el siguiente enlace:

[nbuitrago23/PS1\\_Equipo\\_10](#)