



# CENTRO DE ESTUDIOS ECONÓMICOS

Maestría en Economía 2024–2026

Econometría Aplicada

## Problem Set 2

**PRESENTA:** José Daniel Fuentes García

**PROFESOR:** Raymundo Campos

**LABORATORISTA:** Marcos Ehekatzin

# Índice

<b>Lista de tablas</b>	<b>4</b>
<b>Lista de figuras</b>	<b>6</b>
<b>Instrucciones</b>	<b>11</b>
<b>Problema 1 : ENIGH Y ENOE</b>	<b>11</b>
1.1 . . . . .	11
1.2 . . . . .	12
1.3 . . . . .	12
1.3.1 Resultados ENIGH . . . . .	14
1.3.2 Resultados ENOE . . . . .	17
1.4 . . . . .	20
1.5 . . . . .	20
1.6 . . . . .	21
1.7 . . . . .	22
1.7.1 Resultados ENIGH . . . . .	22
1.7.2 Resultados ENOE . . . . .	23
1.8 . . . . .	24
1.9 . . . . .	51
1.9.1 Resultados ENIGH . . . . .	51
1.9.2 Resultados ENOE . . . . .	55
1.10 . . . . .	57
<b>Problema 2 : IMSS</b>	<b>58</b>
2.1 . . . . .	58
2.2 . . . . .	58
2.3 . . . . .	59
2.4 . . . . .	60
2.5 . . . . .	62
2.6 . . . . .	64
2.7 . . . . .	65
2.8 . . . . .	66
2.8.1 Mapas en  . . . . .	66
2.8.2 Mapas en  . . . . .	69
2.9 . . . . .	71
2.9.1 Mapas en  . . . . .	71
2.9.2 Resultados en  . . . . .	74
2.10 . . . . .	76
2.11 . . . . .	78
2.12 . . . . .	79
<b>Problema 3: Regresión Cuantil</b>	<b>88</b>

3.1 . . . . .	88
3.2 . . . . .	88
3.3 . . . . .	90
3.4 . . . . .	92
3.5 . . . . .	93
3.6 . . . . .	96
3.6.1 Resultados ENIGH . . . . .	97
3.6.2 Resultados ENOE . . . . .	102
<b>Problema 4: Bootstrap</b>	<b>104</b>
4.1 . . . . .	104
a) . . . . .	105
b) . . . . .	106
c) . . . . .	107
d) . . . . .	108
e) . . . . .	108
f) . . . . .	109
4.2 . . . . .	110
4.3 . . . . .	113
<b>Problema 5: Non Parametrics</b>	<b>117</b>
5.1 . . . . .	117
5.2 . . . . .	118
5.3 . . . . .	123
5.4 . . . . .	125
a) . . . . .	125
b) . . . . .	126
c) . . . . .	128
d) . . . . .	131
e) . . . . .	131
i) . . . . .	132
ii) . . . . .	132
iii) . . . . .	135
5.5 . . . . .	137
<b>Problema 6: Imputación</b>	<b>151</b>
6.1 . . . . .	151
6.2 . . . . .	152
6.3 . . . . .	157
6.4 . . . . .	160
6.5 . . . . .	162
6.6 . . . . .	163
6.7 . . . . .	163
<b>Apéndice</b>	<b>167</b>

A: Regresión Cuantil . . . . .	167
B:Estimación no Paramétrica, Trade-off Varianza-Sesgo . . . . .	172
C : Regresiones Cuantiles ENOE . . . . .	173

## **Lista de tablas**

Tabla 1.	Estadísticas descriptivas de la población trabajadora: medias ponderadas por año ENIGH, 1992–2024 . . . . .	14
Tabla 2.	Estadísticas descriptivas de la población trabajadora: desviaciones estándar ponderadas por año ENIGH, 1992–2024 . . . . .	15
Tabla 3.	Estadísticas descriptivas de la población trabajadora: medianas por año ENIGH, 1992–2024 . . . . .	16
Tabla 4.	Estadísticas descriptivas de la población trabajadora: medias ponderadas por año ENOE, 2005–2025 . . . . .	17
Tabla 5.	Desviación estándar de características sociodemográficas y laborales en la ENOE (2005–2025) . . . . .	18
Tabla 6.	Mediana de características sociodemográficas y laborales en la ENOE (2005–2025) . . . . .	19
Tabla 7.	Estadísticas descriptivas: mediana de salarios por año y tipo de medición, ENIGH 1992–2024 . . . . .	32
Tabla 8.	Tasas de crecimiento interanual de salarios reales, ENIGH 1992–2024 . .	33
Tabla 9.	Salario promedio mensual real por grupo de población, ENIGH 1992–2024	34
Tabla 10.	Evolución del salario trimestral real por grupo de población, ENIGH 1992–2024 . . . . .	35
Tabla 11.	Salario mensual por hora promedio según grupo de población, ENIGH 1992–2024 . . . . .	36
Tabla 12.	Salario trimestral por hora promedio según grupo de población, ENIGH 1992–2024 . . . . .	37
Tabla 13.	Crecimiento porcentual del salario mensual real promedio según grupo de población, ENIGH 1992–2024 . . . . .	38
Tabla 14.	Crecimiento porcentual del salario trimestral real promedio según grupo de población, ENIGH 1992–2024 . . . . .	39
Tabla 15.	Crecimiento porcentual del salario trimestral por hora real promedio según grupo de población, ENIGH 1992–2024 . . . . .	40
Tabla 16.	Salario mensual y por hora promedio en la ENOE (2005–2025) . . . .	45
Tabla 17.	Crecimiento interanual del salario mensual y por hora en la ENOE (2005–2025) . . . . .	46
Tabla 18.	Salario mensual promedio por Grupos en la ENOE (2005–2025) . . . .	47
Tabla 19.	Salario por hora promedio por Grupo en la ENOE (2005–2025) . . . .	48
Tabla 20.	Crecimiento interanual del salario mensual por Grupo en la ENOE (2005–2025) . . . . .	49
Tabla 21.	Crecimiento porcentual del salario mensual por hora según grupo poblacional, ENOE (2005–2025) . . . . .	50
Tabla 22.	Proporción de trabajadores en la población por Grupo Poblacional en la ENIGH, 1992–2024 . . . . .	53
Tabla 23.	Proporción de trabajadores en la población con ingreso valido en la ENIGH, 1992–2024 . . . . .	54
Tabla 24.	Proporción de trabajadores en la población por Grupo, ENOE (2005–2025)	55

Tabla 25.	Proporción de trabajadores en la población con ingreso valido en la ENOE, 2005-2025 . . . . .	56
Tabla 26.	Indicadores de ajuste y significancia de los modelos sintéticos de salario (Enero-2015-Julio2025) . . . . .	86
Tabla 27.	Indicadores de ajuste y significancia de los modelos sintéticos de empleo (Enero-2015-Julio2025) . . . . .	86
Tabla 28.	Estimaciones OLS del logaritmo del salario por hora, 2018 . . . . .	93
Tabla 29.	Estimaciones OLS del logaritmo del salario por hora, 2024 . . . . .	93
Tabla 30.	Estimaciones de Mínimos Cuadrados Ordinarios (MCO) del logaritmo del salario por hora, ENIGH 2024 . . . . .	104
Tabla 31.	Comparación de OLS robusto y Bootstrap (100): errores estándar e intervalos de confianza (ENIGH 2024) . . . . .	105
Tabla 32.	Comparación de OLS robusto y Bootstrap (1000): errores estándar e intervalos de confianza (ENIGH 2024) . . . . .	106
Tabla 33.	Comparación de errores estándar por método (ENIGH 2024) . . . . .	107
Tabla 34.	Errores estándar e intervalos de confianza con <b>Bootstrap (100)</b> del <b>25 % de N</b> (ENIGH 2024) . . . . .	108
Tabla 35.	Errores estándar e intervalos de confianza con <b>Bootstrap (1000)</b> del <b>25 % de N</b> (ENIGH 2024) . . . . .	109
Tabla 36.	Errores estándar e intervalos de confianza con <b>bsample</b> 100 iteraciones (ENIGH, 2024) . . . . .	110
Tabla 37.	Errores estándar e intervalos de confianza con <b>bsample</b> 1000 iteraciones (ENIGH, 2024) . . . . .	111
Tabla 38.	Errores estándar e intervalos de confianza con <b>bsample</b> 100 iteraciones y 25 % de la muestra (ENIGH, 2024) . . . . .	111
Tabla 39.	Errores estándar e intervalos de confianza con <b>bsample</b> 1000 iteraciones y 25 % de la muestra (ENIGH, 2024) . . . . .	112
Tabla 40.	Errores estándar e intervalos de confianza con <b>bootstrap</b> 100 iteraciones. (ENOE, 2025 TRIMESTRE II) . . . . .	113
Tabla 41.	Errores estándar e intervalos de confianza con <b>bootstrap</b> 1000 iteraciones. (ENOE, 2025 TRIMESTRE II) . . . . .	113
Tabla 42.	Errores estándar e intervalos de confianza con <b>bootstrap</b> 100 iteraciones y 25 % de la muestra. (ENOE, 2025 TRIMESTRE II) . . . . .	114
Tabla 43.	Errores estándar e intervalos de confianza con <b>bootstrap</b> 1000 iteraciones y 25 % de la muestra. (ENOE, 2025 TRIMESTRE II) . . . . .	114
Tabla 44.	Errores estándar e intervalos de confianza con <b>bsample</b> 100 iteraciones. (ENOE, 2025 TRIMESTRE II) . . . . .	115
Tabla 45.	Errores estándar e intervalos de confianza con <b>bsample</b> 1000 iteraciones. (ENOE, 2025 TRIMESTRE II) . . . . .	115
Tabla 46.	Errores estándar e intervalos de confianza con <b>bsample</b> 100 iteraciones y 25 % de la muestra. (ENOE, 2025 TRIMESTRE II) . . . . .	116
Tabla 47.	Errores estándar e intervalos de confianza con <b>bsample</b> 1000 iteraciones y 25 % de la muestra. (ENOE, 2025 TRIMESTRE II) . . . . .	116

## Listado de figuras

1	Distribución del logaritmo del salario mensual por hora de los trabajadores, por año (ENIGH, 1992–2024) . . . . .	22
2	Distribución del logaritmo del salario trimestral por hora de los trabajadores, por año (ENIGH, 1992–2024) . . . . .	23
3	Distribución del logaritmo del salario mensual por hora de los trabajadores, por año (ENOE, 2005–2025) . . . . .	23
4	Distribución del logaritmo del salario trimestral por hora de los trabajadores, por año (ENOE, 2005–2025) . . . . .	24
5	Bubble Scatter: Salario Real Mensual en Grupo 1 y 7,(ENIGH, 1992–2024) . .	25
6	Bubble Scatter: Salario Real Mensual en Grupo 1 y 5,(ENIGH, 1992–2024) .	26
7	Brechas relativas en el salario real por hora mensual entre grupos de población; Grupo 4=100,(ENIGH, 1992–2024) . . . . .	27
8	Boxplot, Ingreso Mensual Real por Grupo Poblacional,(ENIGH, 2018–2024)	28
9	Tendencias comparadas de salarios reales en México: mensual, trimestral y por hora,(ENIGH, 1992–2024) . . . . .	29
10	Evolución del índice del salario mensual real por grupo de población,(ENIGH, 1992–2024) . . . . .	29
11	Evolución del índice del salario trimestral real por grupo de población,(ENIGH, 1992–2024) . . . . .	30
12	Evolución del índice del salario trimestral por hora real por grupo de población,(ENIGH, 1992–2024) . . . . .	30
13	Evolución del índice del salario mensual por hora real por grupo de población,(ENIGH, 1992–2024) . . . . .	31
14	Evolución del salario real por hora mensual en el primer trimestre: comparación entre Grupo 1 y Grupo 7 (ENOE, Q1, 2005–2025) . . . . .	41
15	Evolución del salario real por hora mensual en Mujeres de baja educación: comparación entre 25–45 y 46–65 años (ENOE, Q1, 2005–2025) . . . . .	41
16	Boxplot, Ingreso Mensual Real por Grupo Poblacional,(ENOE, Q1, 2022–2025)	42
17	Brechas relativas en salario real por hora entre Grupos; Base Grupo 4=100,(ENOE, 2005–2025) . . . . .	43
18	Índices de salarios reales en la ENOE; 2005q1=100,(ENOE, 2005–2025) . .	43
19	Índices de salario mensual por grupo de población en la ENOE; 2005q1=100,(ENOE, 2005–2025) . . . . .	44
20	Índices de salario por hora por grupo de población en la ENOE; 2005q1=100,(ENOE, 2005–2025) . . . . .	44
21	Distribución de trabajadores válidos por grupo poblacional en la ENIGH, 2024) .	51
22	Trabajadores por sexo registrados en el IMSS, 2000–2025 . . . . .	58
23	Promedio y Mediana de ingresos diarios registrados en el IMSS, 2000–2025 .	59
24	Promedio y Mediana por percentil del ingreso diario con datos del IMSS, 2000–2025 . . . . .	60
25	Brecha Salarial (Diario) por Género con datos del IMSS, 2000–2025 . . . . .	61
26	Tamaño de la Brecha Salarial por Género con datos del IMSS, 2000–2025 . .	61

27	Brecha salarial de género estimada por OLS con intervalos de confianza bootstrap (2000–2025) . . . . .	62
28	2 Brecha salarial de género estimada por OLS con intervalos de confianza bootstrap (2000–2025) . . . . .	63
29	Brecha salarial de género en México (2000–2025): promedio anual . . . . .	63
30	Brecha Salarial de Género por Entidad Federativa: Comparación 2020 (Febrero) vs 2025 (Julio) . . . . .	64
31	Crecimiento porcentual del empleo formal por entidad federativa (2012–2025) .	65
32	Crecimiento porcentual del empleo formal por entidad federativa 2017-2025 .	66
33	Crecimiento porcentual del empleo formal por entidad federativa 2010-2025 .	67
34	Crecimiento porcentual del salario por entidad federativa 2017-2025 . . . . .	67
35	Crecimiento porcentual del salario por entidad federativa 2010-2025 . . . . .	68
36	Crecimiento porcentual del empleo formal por entidad federativa 2017-2025 .	69
37	Crecimiento porcentual del empleo formal por entidad federativa 2010-2025 .	69
38	Crecimiento porcentual del salario por entidad federativa 2017-2025 . . . . .	70
39	Crecimiento porcentual del salario por entidad federativa 2010-2025 . . . . .	70
40	Crecimiento porcentual del empleo por entidad federativa Febrero 2020 - Julio 2020 . . . . .	71
41	Crecimiento porcentual del empleo por entidad federativa Febrero 2020 - Julio 2025 . . . . .	72
42	Crecimiento porcentual del salario por entidad federativa Febrero 2020 - Julio 2020 . . . . .	72
43	Crecimiento porcentual del salario por entidad federativa Febrero 2020 - Julio 2025 . . . . .	73
44	Crecimiento porcentual del empleo por entidad federativa Febrero 2020 - Julio 2020 . . . . .	74
45	Crecimiento porcentual del empleo por entidad federativa Febrero 2020 - Julio 2025 . . . . .	74
46	Crecimiento porcentual del salario por entidad federativa Febrero 2020 - Julio 2020 . . . . .	75
47	Crecimiento porcentual del salario por entidad federativa Febrero 2020 - Julio 2025 . . . . .	75
48	Crecimiento porcentual del empleo por municipio 2018-2019 . . . . .	76
49	Crecimiento porcentual del empleo por municipio 2018-2025 . . . . .	77
50	Crecimiento porcentual del salario por municipio 2018-2019 . . . . .	77
51	Crecimiento porcentual del salario por municipio 2018-2025 . . . . .	78
	— Tratamiento	
52	Efectos en la ZLFN en salario promedio (Enero 2015-Julio 2025) . . . . .	80
	--- Sintético	
	— Tratamiento	
53	Placebos de modelos de salario promedio (Enero 2015-Julio 2025) . . . . .	81
54	Placebos RMSPE de modelos de salario promedio (Enero 2015-Julio 2025) . .	82
	— Donantes (placebos)	
	— Tratamiento	
55	Efectos en la ZLFN en empleo (Enero 2015-Julio 2025) . . . . .	83
	--- Sintético	

	Tratamiento	
	Donantes (placebos)	
56	Placebos de modelos de empleo (Enero 2015-Julio 2025)	84
57	Placebos RMSPE de modelos de empleo (Enero 2015-Julio 2025) . . . . .	85
58	Cambio en log salario por hora (2018–2024) por sexo y cuantil . . . . .	91
59	2.Cambio en log salario por hora (2018–2024) por sexo y cuantil . . . . .	91
60	Cambio en log salario por hora, sexo y cuantil (ENOE, 2005–2015) . . . . .	94
61	Cambio en log salario por hora, sexo y cuantil (ENOE, 2015–2025) . . . . .	95
62	Cambio en log salario por hora, sexo y cuantil (ENOE, 2005–2025) . . . . .	95
63	ENIGH: Efecto de la escolaridad en el salario por hora: regresiones cuantiles 2018 vs. 2024 . . . . .	97
64	ENIGH: Efecto del sexo femenino en el salario por hora: regresiones cuantiles 2018 vs. 2024 . . . . .	98
65	ENIGH: Efecto de la escolaridad en el salario por hora: regresiones cuantiles 2018 . . . . .	98
66	ENIGH: Efecto de la escolaridad en el salario por hora: regresiones cuantiles 2024 . . . . .	99
67	ENIGH: Efecto del sexo femenino en el salario por hora: regresiones cuantiles 2018 . . . . .	100
68	ENIGH: Efecto del sexo femenino en el salario por hora: regresiones cuantiles 2024 . . . . .	100
69	ENIGH: Efecto de Perenecer al Sector Rural en el salario por hora: regresiones cuantiles 2018 . . . . .	101
70	ENIGH: Efecto de Perenecer al Sector Rural en el salario por hora: regresiones cuantiles 2024 . . . . .	102
71	ENIGH: Efecto de la escolaridad en el salario por hora: regresiones cuantiles 2015 vs. 2025 . . . . .	102
72	ENIGH: Efecto del sexo femenil en el salario por hora: regresiones cuantiles 2015 vs. 2025 . . . . .	103
73	Estimaciones de densidad kernel del ingreso por hora (ENIGH 1992–2006) .	119
74	Estimaciones de densidad kernel del ingreso por hora (ENIGH 2008–2024) .	120
75	Estimaciones de densidad kernel del logaritmo del ingreso por hora (ENIGH 1992–2006) . . . . .	121
76	Estimaciones de densidad kernel del logaritmo del ingreso por hora (ENIGH 2008–2024) . . . . .	122
77	Comparación de estimaciones de densidad kernel del logaritmo del ingreso por hora bajo diferentes valores de bandwidth (ENIGH 2018) . . . . .	124
78	Comparación de estimaciones de densidad kernel con diferentes funciones kernel (ENIGH 2018) . . . . .	125
79	Relación no paramétrica entre edad y logaritmo del salario por hora en hombres (ENIGH 2018) . . . . .	126
80	Relación no paramétrica entre edad y logaritmo del salario por hora en mujeres (ENIGH 2018) . . . . .	127
81	Relación no paramétrica entre escolaridad y logaritmo del salario por hora en hombres (ENIGH 2018) . . . . .	127

82	Relación no paramétrica entre escolaridad y logaritmo del salario por hora en mujeres (ENIGH 2018) . . . . .	128
83	Relación no paramétrica entre edad y logaritmo del salario por hora en hombres (ENIGH 2024) . . . . .	129
84	Relación no paramétrica entre edad y logaritmo del salario por hora en mujeres (ENIGH 2024) . . . . .	129
85	Relación no paramétrica entre escolaridad y logaritmo del salario por hora en hombres (ENIGH 2024) . . . . .	130
86	Relación no paramétrica entre escolaridad y logaritmo del salario por hora en mujeres (ENIGH 2024) . . . . .	130
87	Relación neta entre edad y logaritmo del salario por hora, controlando por escolaridad, sexo y ruralidad (ENIGH,2018) . . . . .	134
88	Relación neta entre edad y logaritmo del salario por hora, controlando por escolaridad, sexo y ruralidad (ENIGH,2022) . . . . .	135
89	Relación semiparamétrica entre edad e ingreso laboral horario en logaritmos (ENIGH, 2018) . . . . .	136
90	Relación semiparamétrica entre edad e ingreso laboral horario en logaritmos (ENIGH, 2022) . . . . .	136
91	Estimaciones de densidad kernel del ingreso por hora (ENOE, 2005–2013) . .	137
92	Estimaciones de densidad kernel del ingreso por hora (ENOE, 2014–2022) . .	138
93	Estimaciones de densidad kernel del ingreso por hora (ENOE, 2023–2025) . .	139
94	Estimaciones de densidad kernel del logaritmo del ingreso por hora (ENOE, 2005–2013) . . . . .	140
95	Estimaciones de densidad kernel del logaritmo del ingreso por hora (ENOE, 2014–2022) . . . . .	141
96	Estimaciones de densidad kernel del logaritmo del ingreso por hora (ENOE, 2023–2025) . . . . .	142
97	Comparación de estimaciones de densidad kernel del logaritmo del ingreso por hora bajo diferentes valores de bandwidth (ENOE 2025) . . . . .	142
98	Comparación de estimaciones de densidad kernel con diferentes funciones kernel (ENOE, 2025) . . . . .	143
99	Relación no paramétrica entre edad y logaritmo del salario por hora en hombres (ENOE 2018) . . . . .	144
100	Relación no paramétrica entre edad y logaritmo del salario por hora en mujeres (ENOE 2018) . . . . .	144
101	Relación no paramétrica entre esolaridad y logaritmo del salario por hora en hombres (ENOE 2018) . . . . .	145
102	Relación no paramétrica entre esolaridad y logaritmo del salario por hora en mujeres (ENOE 2018) . . . . .	145
103	Relación no paramétrica entre edad y logaritmo del salario por hora en hombres (ENOE 2024) . . . . .	146
104	Relación no paramétrica entre edad y logaritmo del salario por hora en mujeres (ENOE 2024) . . . . .	146
105	Relación no paramétrica entre escolaridad y logaritmo del salario por hora en hombres (ENOE 2024) . . . . .	147

106	Relación no paramétrica entre escolaridad y logaritmo del salario por hora en mujeres (ENOE 2024) . . . . .	147
107	Relación neta entre edad y logaritmo del salario por hora, controlando por escolaridad, sexo y ruralidad (ENOE,2018) . . . . .	148
108	Relación neta entre edad y logaritmo del salario por hora, controlando por escolaridad, sexo y ruralidad (ENOE,2022) . . . . .	148
109	Relación semiparamétrica entre edad e ingreso laboral horario en logaritmos (ENOE, 2018) . . . . .	149
110	Relación semiparamétrica entre edad e ingreso laboral horario en logaritmos (ENOE, 2022) . . . . .	149
111	Evolución trimestral del porcentaje de trabajadores que no reportan ingreso (ENOE, 2005-2025q2) . . . . .	154
112	Evolución trimestral del porcentaje de trabajadores que no responde ingresos y rango de salario (ENOE, 2005-2025q2) . . . . .	154
113	Evolución trimestral del porcentaje de trabajadores que no reportan ingreso segun situacion formal (ENOE, 2005-2025q2) . . . . .	155
114	Evolución trimestral del porcentaje de trabajadores que no reportan ingreso por sexo (ENOE, 2005-2025q2) . . . . .	155
115	Evolución trimestral del porcentaje de trabajadores que no reportan ingreso por nivel educativo (ENOE, 2005-2025q2) . . . . .	156
116	Evolución trimestral del porcentaje de trabajadores que no reportan ingreso por sector (ENOE, 2005-2025q2) . . . . .	156
117	Evolución De Poblacion en Pobreza Multidimensional y Pobreza Laboral,(ENIGH y ENOE, 2016-2024) . . . . .	157
118	Relación entre Pobreza Multidimensional y Pobreza Laboral,(ENIGH y ENOE, 2016-2024) . . . . .	158
119	Evolución De Poblacion en Pobreza Laboral (ENOE, 2005-2025q2) . . . . .	158
120	Evolución De Poblacion en Pobreza Multidimensional, INEGI (ENIGH, 2016-2024) . . . . .	159
121	Porcentaje de la Poblacion en Pobreza Laboral,(ENOE, 2025 TRIMSTRE II)	159
122	Porcentaje de la Poblacion en Pobreza Multidimensional,(ENIGH, 2024) . .	160
123	Ingreso Reportado vs Hotdeck vs iPolate (ENOE,2005-2025q2) . . . . .	161
124	% de la Población en ITLP: Reportado vs Hotdeck vs iPolate (ENOE,2005-2025q2) . . . . .	162
125	Correlación entre ambios en trabajadores con ingresos inválidos y cambios la tasa de participación laboral (ENOE,2005-2025q2) . . . . .	164
126	Correlación entre trabajadores con ingresos inválidos y la tasa de participación laboral (ENOE,2005-2025q2) . . . . .	165
127	Relación de cambios porcentuales entre de pobreza laboral y número de asalariados (2018q2,2025q2) . . . . .	165
128	Relación de cambios porcentuales entre de pobreza laboral e ingreso promedio diario (2018q2,2025q2) . . . . .	166
129	Regresiones Cuantiles (ENOE: 2005,2015,2025) . . . . .	173

## Instrucciones

---

**Fecha límite:** Lunes **29 de septiembre**, antes de clase, en **Teams**.

El objetivo del *problem set* es que el estudiante analice los datos de la **ENIGH**, **ENOE**, **IMSS** y pueda realizar estadísticas descriptivas con esos datos. Se pretende que mejore sustancialmente su uso y conocimiento de microdatos como de **Stata** y **R**.

Comandos que se utilizarán para buscar en ayuda: `merge`, `collapse`, `append`, `table`, `sum`. Asimismo analiza la discusión en *Stata* sobre los pesos para generalizar los resultados a toda la población. En particular utilizaremos `fweight` y `aweight`.

Recomiendo que tengas una carpeta dentro de *Econometría* o algo así, y organices tus programas de *Stata* ahí. Por ejemplo, yo siempre que empiezo un proyecto tengo 4 carpetas en ese proyecto, una llamada `data` para los datos, otra llamada `prog` para los programas, otra llamada `log` para los log files, y otra para papers o algo así. Cuando escribas tu programa sólo tienes que llamar las bases dentro de esa carpeta, etc.

Asegúrate que tus gráficas y tablas sean presentables como si las fueras a incluir en un *paper*. Es decir, ponle nombre a los ejes, utiliza diferentes tipos de líneas, etc. **Gráficas y tablas que no se vean bonitas y profesionales serán castigadas.**

## Problema 1: ENIGH Y ENOE

---

El **objetivo** es tener características de individuos y salarios para diferentes puntos en el tiempo.

### 1.1

Las bases que utilizaremos son todas las **ENIGH** hasta 2024, y todas las **ENOE** de todos los trimestres hasta 2025 (*sin ETOE, si tienes problema con el peso, puedes quedarte con el 2o trimestre de cada año*).

Es muy importante que las bases de datos queden limpias y uniformes en el `#` de observaciones con todas las variables utilizadas, pues las bases construidas aquí serán utilizadas en los siguientes problemas.

**RESPUESTA:** Requisito cumplido ✓

## 1.2

**ENIGH:** De la tabla de ingresos obtén el ingreso del mes pasado por trabajo y el ingreso trimestral normalizado y pégalo a la base de población. En la base de ingreso tienes que escoger el ingreso del mes pasado (*fíjate en la codificación de las variables*), y luego hacer un `collapse` para el folio del individuo. Después tienes que hacer un `merge xxx, by(folio)` con la base de población. Esto lo tienes que hacer para cada año, y después juntar los años con el comando `append` (*no se te olvide crear una variable gen year=1996 para el año 1996, y otros años*).

El ingreso por trabajo debe de ser homogéneo a través de los años. Para ayuda leer el artículo de **Campos, Lustig y Santillán (2014)** en *Estudios Económicos*.

Al final de este ejercicio, debes tener una base a nivel individual con las muestras de todos los años, similar para la **ENOE**. **ENIGH:** 1992–2024. **ENOE:** 2005–2025 (*no usar ETOE*).

**RESPUESTA:** Requisito cumplido ✓

## 1.3

**ENIGH y ENOE:** Las variables importantes son edad, educación (*años y por nivel terminado*), ingreso del mes pasado, ingreso trimestral normalizado, horas trabajadas, sexo, rural. Asegúrate de tener el ingreso en términos reales, digamos Enero de 2025 (*busca en el Banco de México el INPC para que puedas tener el ingreso constante*). Restringe las observaciones a individuos entre 25–65 años de edad. (*ENOE solo se tiene un ingreso laboral*).

Calcula un `sum`, `detail` de cada una de las variables para cada año o bien una tabla con diferentes estadísticos para que estés seguro que la limpieza de los datos es correcta, y que no tienes *missing values*. Es decir, son las mismas observaciones para todas las variables.

Asegúrate que eso se cumple, y comenta los pasos de limpieza así como las estadísticas finales. También puedes checar el comando `tabstat` o `table`. Averigua comandos `putexcel` o `estout` para exportar a Excel.

**RESPUESTA:**

### 1.3.1 Resultados ENIGH

**Tabla 1.** Estadísticas descriptivas de la población trabajadora: medias ponderadas por año ENIGH, 1992–2024

Año	Edad	Nivel de educación	Salario mensual real	Salario trimestral real	Horas de trabajo	Sexo (1=Mujer)	Urbano o rural
1992	39.82	1.46	9248.75	26441.75	28.57	0.52	0.23
1994	40.31	1.50	10757.46	30784.73	29.64	0.52	0.23
1996	39.80	1.60	7259.23	20461.67	30.61	0.53	0.22
1998	40.37	1.67	7635.50	21493.41	30.64	0.53	0.22
2000	40.61	1.76	8565.84	24825.24	30.86	0.52	0.22
2002	41.01	1.79	8598.63	24967.13	31.30	0.53	0.22
2004	40.92	1.91	8701.58	24923.24	46.96	0.54	0.21
2005	41.32	1.95	8694.88	25036.67	32.37	0.53	0.21
2006	40.95	1.99	9112.83	26298.27	33.50	0.54	0.21
2008	41.58	2.00	8606.46	24786.43	47.24	0.53	0.21
2010	41.82	2.07	8306.42	23735.00	47.87	0.53	0.20
2012	42.20	2.11	8047.72	22418.51	48.68	0.53	0.21
2014	42.18	2.19	8367.46	23341.71	48.53	0.53	0.21
2016	42.26	2.25	8186.53	22906.30	48.73	0.53	0.21
2018	42.57	2.31	8104.86	22761.21	48.79	0.53	0.22
2020	43.00	2.38	7816.93	21260.93	47.51	0.53	0.21
2022	43.28	2.43	8336.76	23443.25	48.41	0.54	0.23
2024	43.44	2.51	9187.50	26062.26	47.69	0.54	0.20

Fuente: Elaboración propia con datos de la ENIGH. Proceso disponible en:  danifuentesga

**Tabla 2.** Estadísticas descriptivas de la población trabajadora: desviaciones estándar ponderadas por año ENIGH, 1992–2024

Año	Edad	Nivel de educación	Salario mensual real	Salario trimestral real	Horas de trabajo	Sexo (1=Mujer)	Urbano o rural
1992	10.92	1.16	10508.04	30427.85	26.82	0.50	0.42
1994	11.19	1.19	15257.41	44471.11	26.92	0.50	0.42
1996	10.96	1.19	8713.74	25405.94	26.78	0.50	0.41
1998	11.05	1.18	9460.85	26954.00	26.47	0.50	0.41
2000	11.04	1.20	11444.60	33614.74	25.78	0.50	0.42
2002	11.02	1.21	9825.27	29108.92	26.05	0.50	0.42
2004	11.01	1.22	10182.87	29864.54	17.29	0.50	0.41
2005	11.11	1.23	12429.51	36842.83	25.71	0.50	0.41
2006	10.95	1.22	10763.83	31788.83	25.82	0.50	0.41
2008	11.03	1.22	9855.72	28370.80	16.24	0.50	0.40
2010	11.04	1.22	9863.94	29599.27	17.43	0.50	0.40
2012	11.26	1.21	9328.81	26244.92	19.21	0.50	0.41
2014	11.13	1.21	16950.18	35400.42	17.66	0.50	0.41
2016	11.16	1.19	9043.27	26269.81	17.89	0.50	0.41
2018	11.33	1.19	8843.66	25934.17	17.60	0.50	0.42
2020	11.44	1.19	7112.98	20632.83	17.65	0.50	0.41
2022	11.49	1.18	8428.22	24226.03	16.91	0.50	0.42
2024	11.48	1.18	8355.04	23773.86	15.88	0.50	0.40

Fuente: Elaboración propia con datos de la ENIGH. Proceso disponible en:  danifuentesga

**Tabla 3.** Estadísticas descriptivas de la población trabajadora: medianas por año ENIGH, 1992–2024

Año	Edad	Nivel de educación	Salario mensual real	Salario trimestral real	Horas de trabajo	Sexo (1=Mujer)	Urbano o rural
1992	38.00	1.00	6543.92	18404.78	35.00	1.00	0.00
1994	38.00	1.00	6967.75	19398.23	35.00	1.00	0.00
1996	38.00	1.00	4801.11	13827.21	36.00	1.00	0.00
1998	38.00	1.00	5163.69	14557.22	39.00	1.00	0.00
2000	39.00	1.00	5809.53	16783.08	40.00	1.00	0.00
2002	40.00	1.00	5777.93	17333.78	40.00	1.00	0.00
2004	39.00	2.00	6158.98	17294.42	48.00	1.00	0.00
2005	40.00	2.00	6092.00	17816.67	40.00	1.00	0.00
2006	39.00	2.00	6531.82	18109.80	40.00	1.00	0.00
2008	40.00	2.00	5976.31	18075.96	48.00	1.00	0.00
2010	40.00	2.00	5721.60	16481.49	48.00	1.00	0.00
2012	41.00	2.00	5770.10	15906.35	48.00	1.00	0.00
2014	41.00	2.00	5625.02	16425.53	48.00	1.00	0.00
2016	41.00	2.00	5991.48	16680.10	48.00	1.00	0.00
2018	42.00	2.00	5984.71	17563.83	48.00	1.00	0.00
2020	42.00	2.00	6050.45	16480.42	48.00	1.00	0.00
2022	43.00	2.00	6530.21	18685.63	48.00	1.00	0.00
2024	43.00	2.00	7386.37	21027.05	48.00	1.00	0.00

Fuente: Elaboración propia con datos de la ENIGH. Proceso disponible en:  danifuentesga

### 1.3.2 Resultados ENOE

**Tabla 4.** Estadísticas descriptivas de la población trabajadora: medias ponderadas por año ENOE, 2005-2025

Año	Edad	Nivel de educación	Ingreso mensual	Ingreso por hora	Horas de trabajo	Sexo (1=Mujer)	Urbano o rural
2005	41.29	1.92	4615.82	26.89	28.04	0.54	0.21
2006	41.31	1.96	4888.30	28.72	28.34	0.54	0.21
2007	41.48	1.98	5104.42	30.43	28.13	0.54	0.21
2008	41.56	2.00	5199.52	30.83	28.23	0.54	0.20
2009	41.72	2.05	5123.28	30.86	27.67	0.53	0.20
2010	41.76	2.08	5175.36	31.16	27.76	0.53	0.20
2011	41.84	2.11	5265.11	31.56	28.00	0.53	0.20
2012	42.06	2.14	5414.93	32.77	28.40	0.53	0.20
2013	42.11	2.17	5561.28	33.59	28.42	0.53	0.20
2014	42.20	2.19	5517.56	33.21	28.26	0.53	0.20
2015	42.28	2.21	5742.24	34.68	28.52	0.53	0.21
2016	42.38	2.24	5982.96	35.86	28.77	0.53	0.21
2017	42.50	2.27	6238.13	37.44	28.83	0.53	0.21
2018	42.64	2.32	6544.06	39.14	29.01	0.53	0.21
2019	42.85	2.34	6860.11	41.40	29.05	0.53	0.21
2020	42.97	2.39	7280.57	46.16	26.91	0.53	0.21
2021	43.16	2.42	7774.71	49.41	27.88	0.53	0.21
2022	43.25	2.44	8359.29	51.12	29.03	0.53	0.21
2023	43.39	2.46	9252.80	56.77	29.30	0.54	0.21
2024	43.52	2.49	10231.13	63.31	29.09	0.54	0.20
2025	43.59	2.52	10842.78	69.09	28.61	0.54	0.20

**Tabla 5.** Desviación estándar de características sociodemográficas y laborales en la ENOE (2005–2025)

Año	Edad	Nivel de educación	Ingreso mensual	Ingreso por hora	Horas de trabajo	Sexo (1=Mujer)	Urbano o rural
2005	11.02	1.22	6326.68	40.20	25.12	0.50	0.41
2006	10.97	1.22	5699.02	39.84	24.99	0.50	0.41
2007	10.97	1.22	6089.03	45.69	24.90	0.50	0.41
2008	10.98	1.21	6032.74	42.68	24.94	0.50	0.40
2009	11.03	1.22	5623.56	41.94	24.92	0.50	0.40
2010	11.06	1.21	5362.63	39.10	24.92	0.50	0.40
2011	11.09	1.21	5298.57	42.85	24.84	0.50	0.40
2012	11.11	1.21	5784.64	57.19	24.84	0.50	0.40
2013	11.09	1.21	7320.97	56.96	24.80	0.50	0.40
2014	11.09	1.21	5312.18	41.91	24.64	0.50	0.40
2015	11.18	1.20	5825.40	66.28	24.74	0.50	0.40
2016	11.22	1.20	5849.65	44.18	24.71	0.50	0.40
2017	11.24	1.20	5772.82	45.82	24.60	0.50	0.40
2018	11.29	1.19	6003.93	47.54	24.54	0.50	0.41
2019	11.34	1.19	6165.34	54.61	24.50	0.50	0.41
2020	11.38	1.19	6895.19	62.93	24.48	0.50	0.41
2021	11.40	1.19	7125.03	67.92	24.57	0.50	0.41
2022	11.43	1.19	7359.06	75.06	24.37	0.50	0.41
2023	11.50	1.19	8244.44	78.96	24.19	0.50	0.41
2024	11.52	1.19	9160.06	78.11	24.06	0.50	0.40
2025	11.50	1.19	9592.51	109.91	23.95	0.50	0.40

Fuente: Elaboración propia con datos de la ENOE. Proceso disponible en:  danifuentesga

**Tabla 6.** Mediana de características sociodemográficas y laborales en la ENOE (2005–2025)

Año	Edad	Nivel de educación	Ingreso mensual	Ingreso por hora	Horas de trabajo	Sexo (1=Mujer)	Urbano o rural
2005	40.00	2.00	3440.00	17.83	32.00	1.00	0.00
2006	40.00	2.00	3600.00	19.38	32.00	1.00	0.00
2007	40.00	2.00	3870.00	20.00	32.00	1.00	0.00
2008	40.00	2.00	4000.00	20.83	32.00	1.00	0.00
2009	40.00	2.00	4000.00	20.83	30.00	1.00	0.00
2010	40.00	2.00	4000.00	21.28	30.00	1.00	0.00
2011	41.00	2.00	4300.00	22.22	32.00	1.00	0.00
2012	41.00	2.00	4300.00	22.73	32.00	1.00	0.00
2013	41.00	2.00	4300.00	23.26	34.00	1.00	0.00
2014	41.00	2.00	4300.00	23.81	34.00	1.00	0.00
2015	41.00	2.00	4400.00	24.92	35.00	1.00	0.00
2016	41.00	2.00	4838.00	25.00	35.00	1.00	0.00
2017	42.00	2.00	5160.00	26.79	35.00	1.00	0.00
2018	42.00	2.00	5160.00	28.57	36.00	1.00	0.00
2019	42.00	2.00	5590.00	30.00	36.00	1.00	0.00
2020	42.00	2.00	6000.00	32.56	30.00	1.00	0.00
2021	43.00	2.00	6450.00	34.38	32.00	1.00	0.00
2022	43.00	2.00	6880.00	37.50	36.00	1.00	0.00
2023	43.00	2.00	7740.00	41.67	36.00	1.00	0.00
2024	43.00	2.00	8600.00	46.51	36.00	1.00	0.00
2025	43.00	2.00	8600.00	50.00	35.00	1.00	0.00

Fuente: Elaboración propia con datos de la ENOE. Proceso disponible en:  danifuentesga

## 1.4

Crea **8 grupos de observaciones**. Crea grupos para **sexo-edad-educación**. Donde grupo de edad es definida como **25–45** y **46–65**, y grupo de educación es definida como *menos de preparatoria o más o igual que preparatoria*.

**RESPUESTA:**

Los grupos de población se definieron según edad, sexo y nivel educativo:

- **Grupo 1** = Mujeres 25–45 años, < Preparatoria;
- **Grupo 2** = Hombres 25–45 años, < Preparatoria;
- **Grupo 3** = Mujeres 25–45 años,  $\geq$  Preparatoria;
- **Grupo 4** = Hombres 25–45 años,  $\geq$  Preparatoria;
- **Grupo 5** = Mujeres 46–65 años, < Preparatoria;
- **Grupo 6** = Hombres 46–65 años, < Preparatoria;
- **Grupo 7** = Mujeres 46–65 años,  $\geq$  Preparatoria;
- **Grupo 8** = Hombres 46–65 años,  $\geq$  Preparatoria.

## 1.5

Si el salario es igual a **999999 (8)** quiere decir que ese salario no es válido, asegúrate de que ese salario sea convertido a “*missing value*” (*antes de ser cambiado a real, los 0's tampoco los usamos*).

Define una variable llamada **trabajo** que sea *dummy variable* para los trabajadores, donde trabajador es definido como aquel con un salario válido. Cambia el salario a pesos reales de **Enero 2024**.

**RESPUESTA:** Requisito cumplido ✓

## 1.6

Crea una variable de **salario por hora**. Es decir, divide el salario por horas trabajadas en la semana multiplicadas por **4.33** para asegurarte que el salario es por hora.

Haz un *censoring* de los datos, es decir, todo salario por hora menor a 1 cámbialo a valor 1, mientras que todos los salarios por hora mayores a **\$5000**, restríngelos a **\$5000**. Asegúrate de no cambiar *missing values* o valores cero.

**¿Cuántas observaciones cambias, a qué se debe?**

### **RESPUESTA:**

#### **Construcción de la variable salario por hora**

Primero, creamos una variable de **salario por hora**, dividiendo el ingreso entre las horas trabajadas por semana, multiplicadas por 4.33. Esto asegura que el valor esté expresado por hora:

```
gen salario_trimestral_hora = ingreso_trimestral_real / (htrab * 4.33 * 3)  
gen salario_mensual_hora = ingreso_mensual_real / (htrab * 4.33)
```

Esto generó muchos *missing values* (**579,680** y **583,709** respectivamente), debido a datos faltantes en **htrab** o **ingresos**.

Luego, censuramos valores extremos para evitar errores o valores poco realistas:

- Si el salario por hora es menor o igual a 1, lo cambiamos a 1.
- Si es mayor o igual a 5000, lo limitamos a 5000.
- **salario\_mensual\_hora**: **807** fueron cambiadas a 1 y **7** a 5000.
- **salario\_trimestral\_hora**: **4,728** a 1 y **6** a 5000.

Estos cambios se deben a registros con ingresos extremadamente bajos o altos, posiblemente errores de captura o unidades incorrectas. Se tuvo cuidado de no modificar *missing values* ni valores igual a cero.

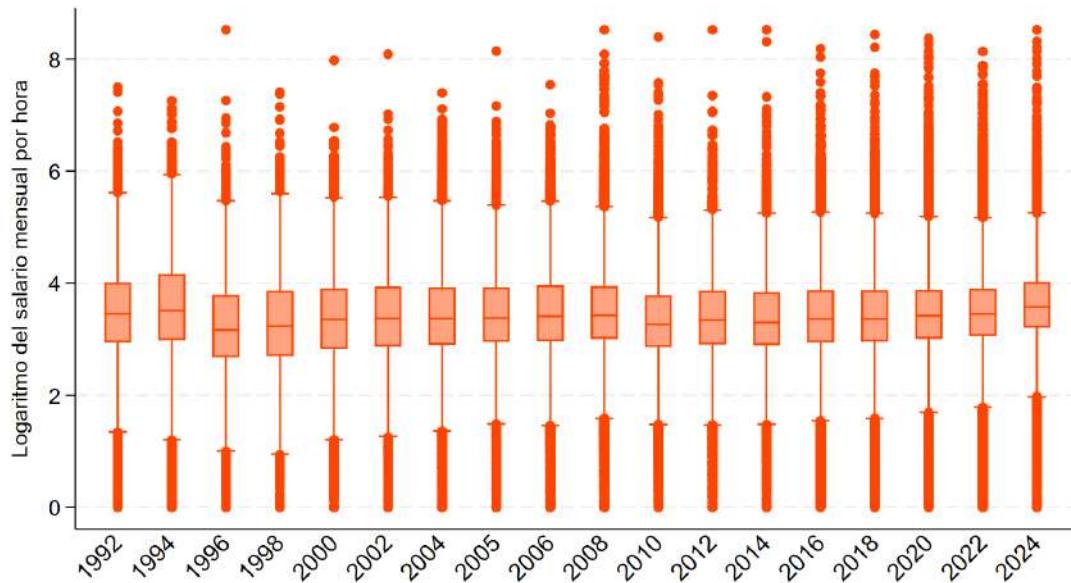
## 1.7

Analiza la dispersión de los datos usando figuras de *box plot*. Usa el **logaritmo del salario por hora**. Pon en eje  $x$  los años. Entender significancia y valores en el *boxplot*.

### RESPUESTA:

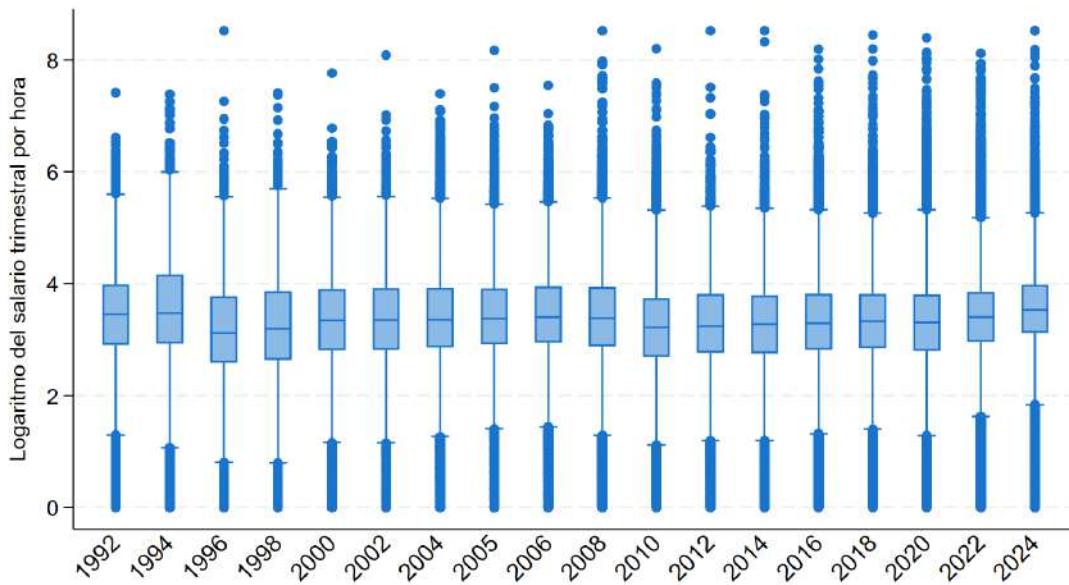
#### 1.7.1 Resultados ENIGH

Figura 1: Distribución del logaritmo del salario mensual por hora de los trabajadores, por año (ENIGH, 1992–2024)



Fuente: Elaboración propia con datos de la ENIGH. Proceso disponible en:  danifuentesga

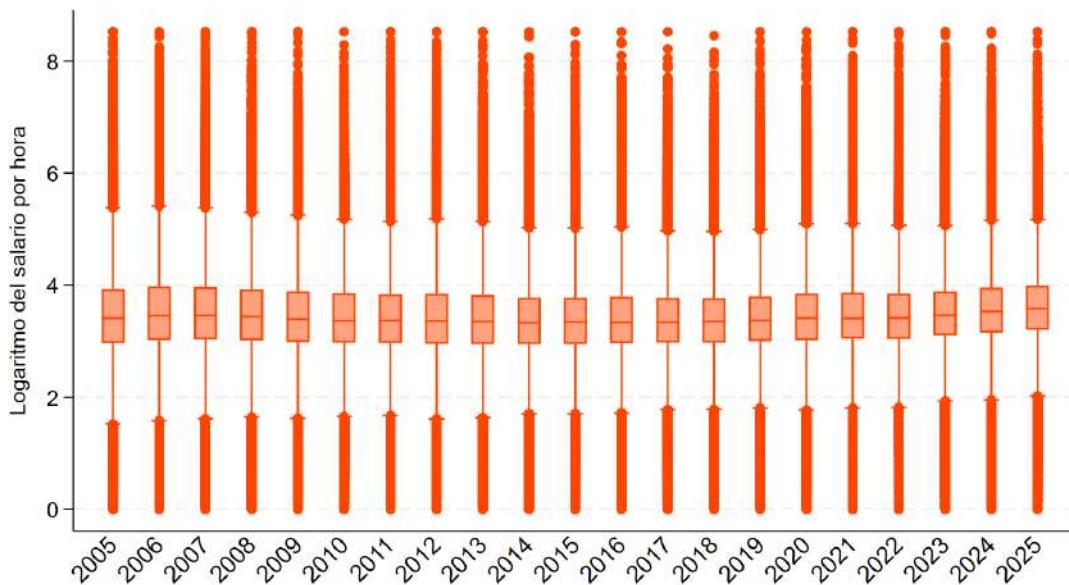
Figura 2: Distribución del logaritmo del salario trimestral por hora de los trabajadores, por año (ENIGH, 1992–2024)



Fuente: Elaboración propia con datos de la ENIGH. Proceso disponible en:  danifuentesga

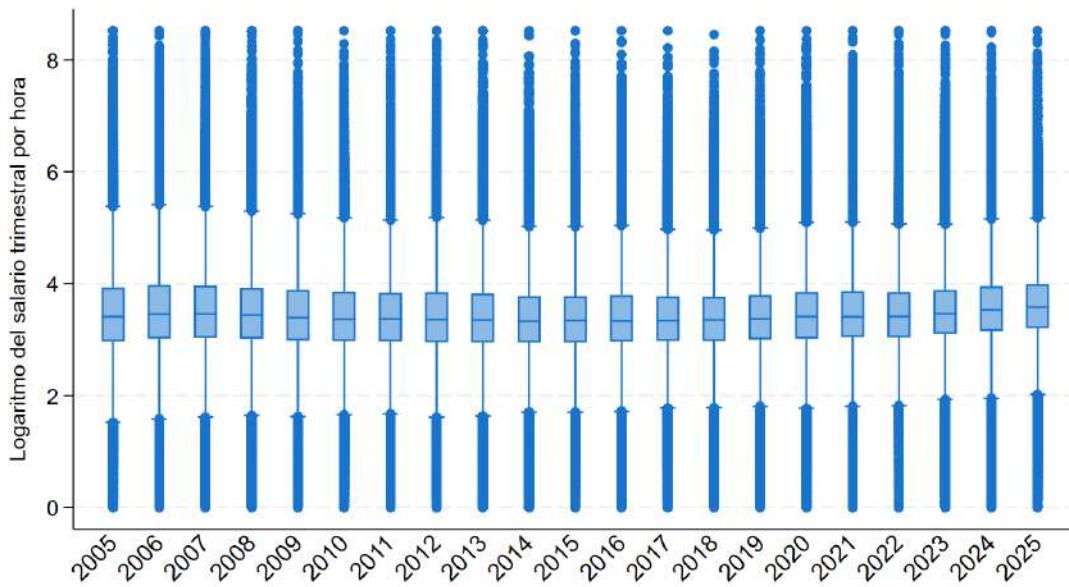
### 1.7.2 Resultados ENOE

Figura 3: Distribución del logaritmo del salario mensual por hora de los trabajadores, por año (ENOE, 2005-2025)



Fuente: Elaboración propia con datos de la ENOE. Proceso disponible en:  danifuentesga

Figura 4: Distribución del logaritmo del salario trimestral por hora de los trabajadores, por año (ENOE,2005-2025)



Fuente: Elaboración propia con datos de la ENOE. Proceso disponible en:  danifuentesga

## 1.8

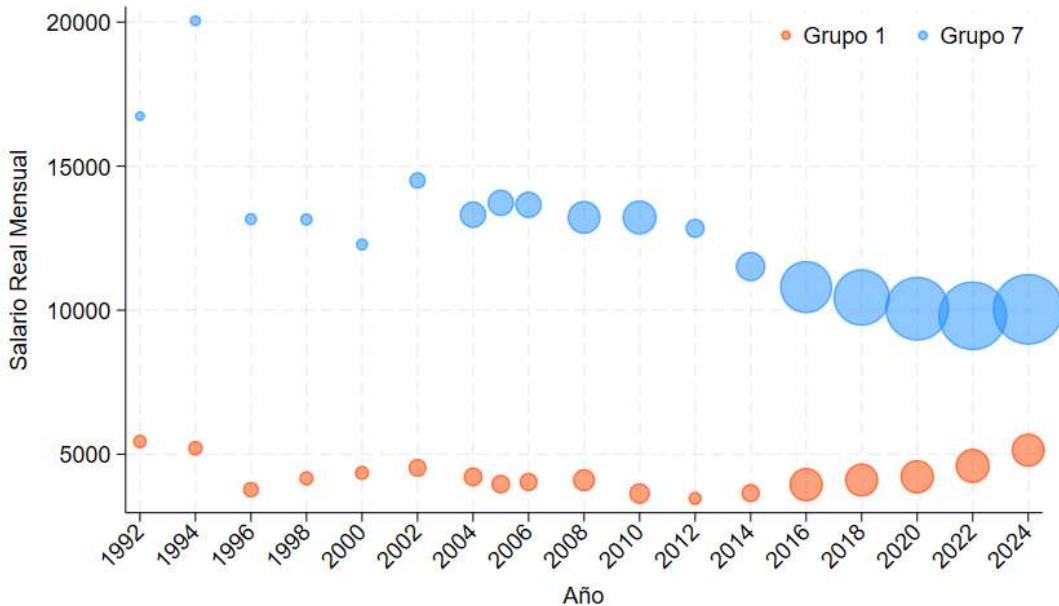
Analiza cómo cambia el salario y el salario por hora para la población y para cada grupo (*asegúrate de utilizar los pesos factor que incluye la ENIGH y ENOE, averigua en el help las propiedades de fweight y aweight*) **ÚNICAMENTE** para la población con salario por hora válido.

De ahora en adelante cuando hablemos de cálculos de salario o salario por hora me refiero **ÚNICAMENTE** a aquellos con ingreso válido (*es decir ingreso positivo, todos los ingresos ceros, missing, de aquellos que no trabajan o no reciben pago NO se incluyen*).

Utiliza el comando `table` y realiza gráficas al respecto. Discute qué grupo ha sido más afectado o beneficiado. Utiliza el comando de pesos `aweight`, utiliza el comando `fweight` y fíjate si lo puedes utilizar. Analiza las razones de porqué no lo puedes utilizar dado sea el caso. Transforma los pesos a números enteros para que puedas utilizar el comando `fweight`. Compara tus resultados.

## RESPUESTA:

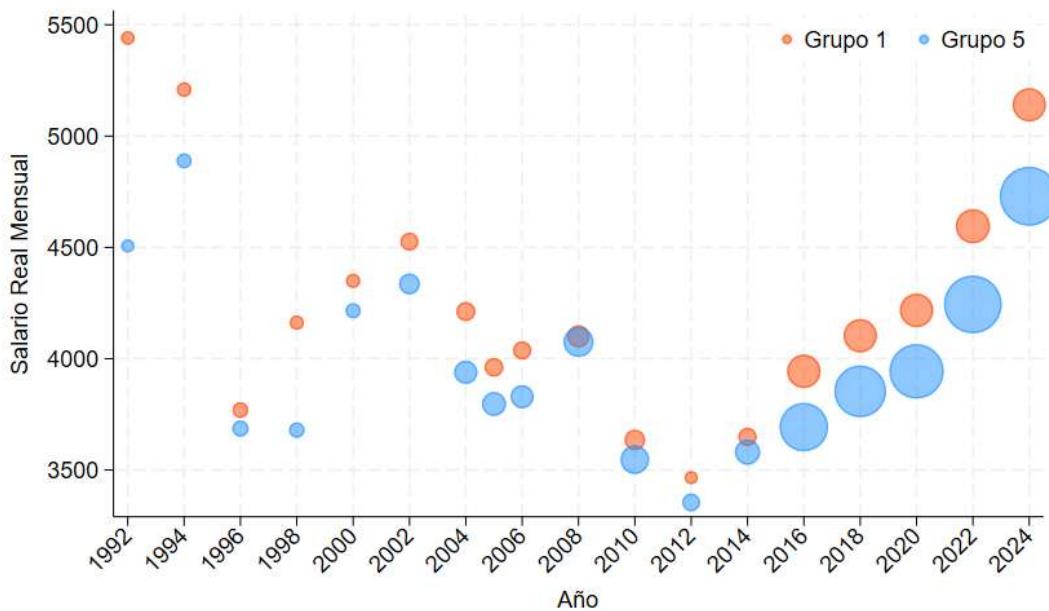
Figura 5: Bubble Scatter: Salario Real Mensual en Grupo 1 y 7,(ENIGH, 1992–2024)



Fuente: Elaboración propia con datos de la ENIGH. Proceso disponible en:  danifuentesga

El gráfico compara la evolución del **salario mensual real** de dos grupos de mujeres entre 1992 y 2024: aquellas de 25 a 65 años con **menos de preparatoria concluida** (grupo 1) y aquellas con **al menos preparatoria completa o mayor educación** (grupo 7). Se observa que el **grupo 1** ha mantenido salarios bajos, en un rango de 3,500 a 5,500 pesos, con una ligera recuperación en los años más recientes. En contraste, el **grupo 7** presenta salarios considerablemente más altos, entre 12,000 y 20,000 pesos, aunque con una tendencia descendente después de 2006 y cierta estabilización hacia 2020–2024. El tamaño creciente de las burbujas indica que en ambos grupos ha aumentado el **número de trabajadoras válidas** registradas en la ENIGH, lo cual refleja una **mayor participación laboral femenina**, pero también la persistencia de una marcada **brecha salarial asociada al nivel educativo**.

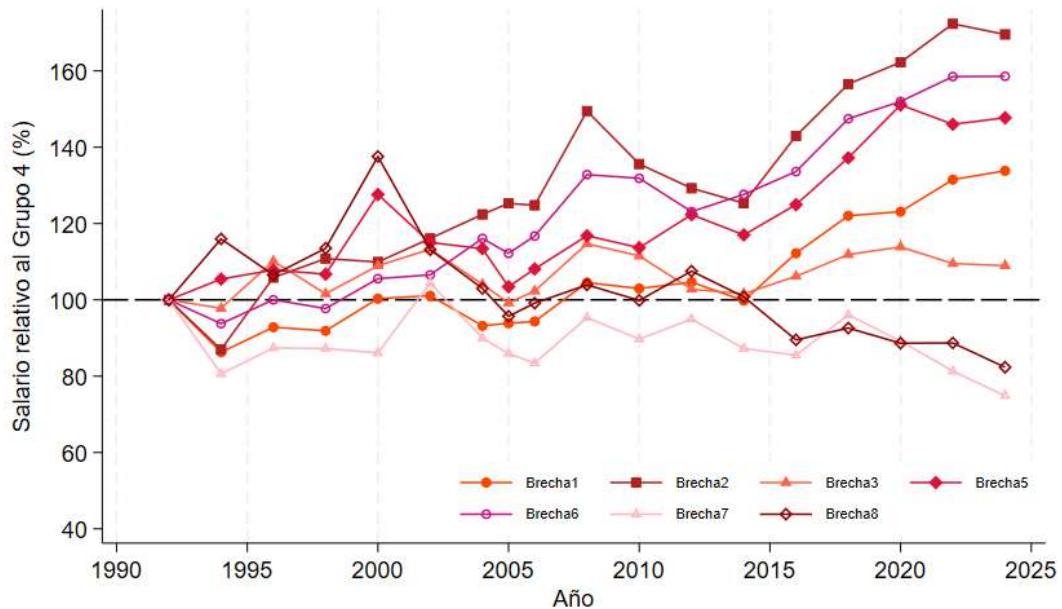
Figura 6: Bubble Scatter: Salario Real Mensual en Grupo 1 y 5,(ENIGH, 1992–2024)



Fuente: Elaboración propia con datos de la ENIGH. Proceso disponible en:  danifuentesga

El gráfico muestra la evolución del **salario mensual real promedio** para mujeres de 25–45 años sin preparatoria concluida (grupo 1) y mujeres de 46–65 años con el mismo nivel educativo (grupo 5), entre 1992 y 2024. Si bien ambos grupos comparten una trayectoria descendente hasta inicios de la década de 2010 y una **recuperación salarial posterior**, la brecha se mantiene: las mujeres mayores (grupo 5) presentan ingresos ligeramente inferiores en casi todo el periodo. El **tamaño creciente de las burbujas** refleja un aumento en el número de trabajadoras en ambos grupos, pero los beneficios en salario son más limitados para las de mayor edad, lo que evidencia una **desventaja acumulada en el mercado laboral** para las mujeres de 46–65 años sin educación media superior.

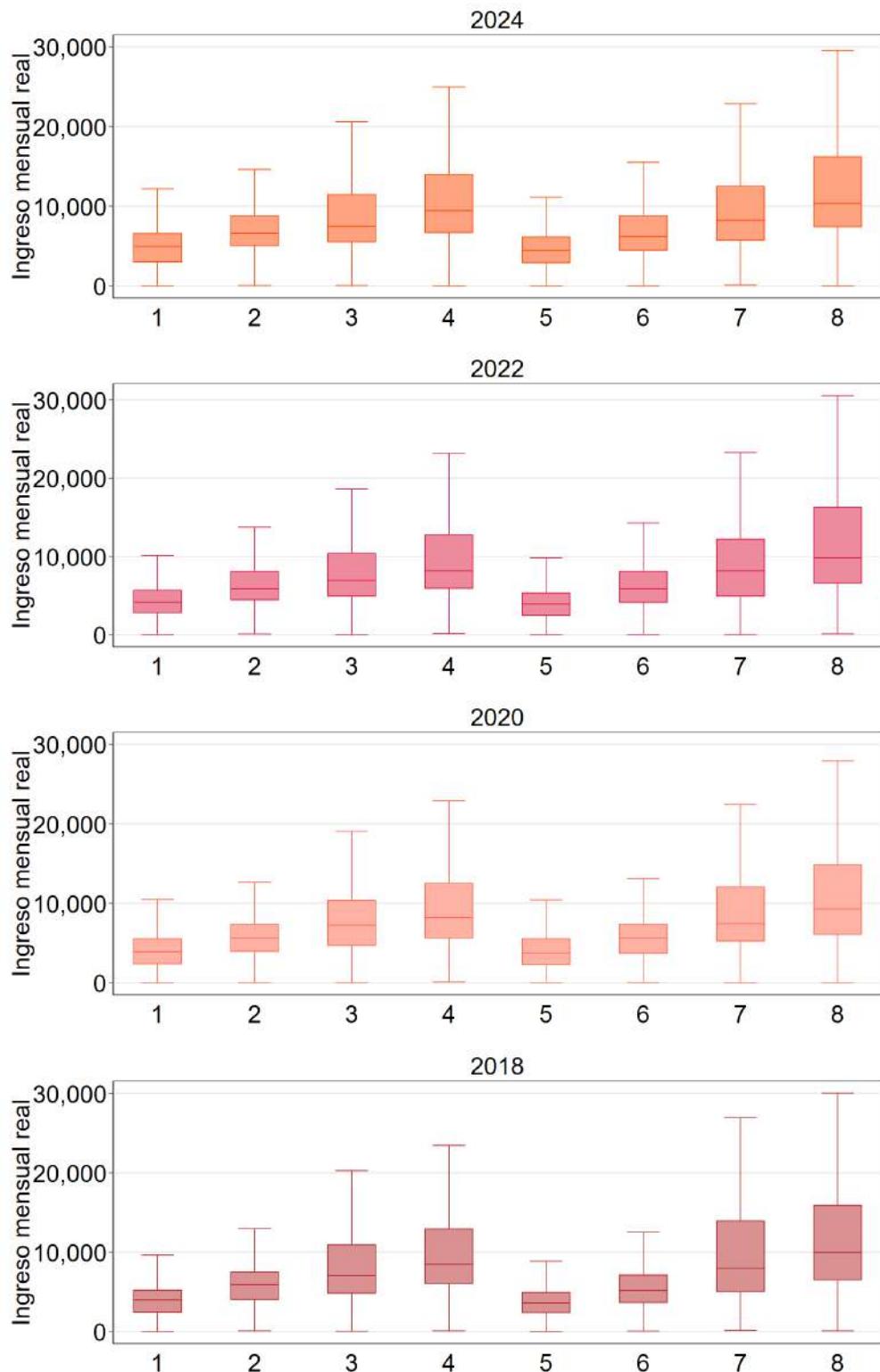
Figura 7: Brechas relativas en el salario real por hora mensual entre grupos de población; Grupo 4=100,(ENIGH, 1992–2024)



Fuente: Elaboración propia con datos de la ENIGH. Proceso disponible en:  danifuentesga

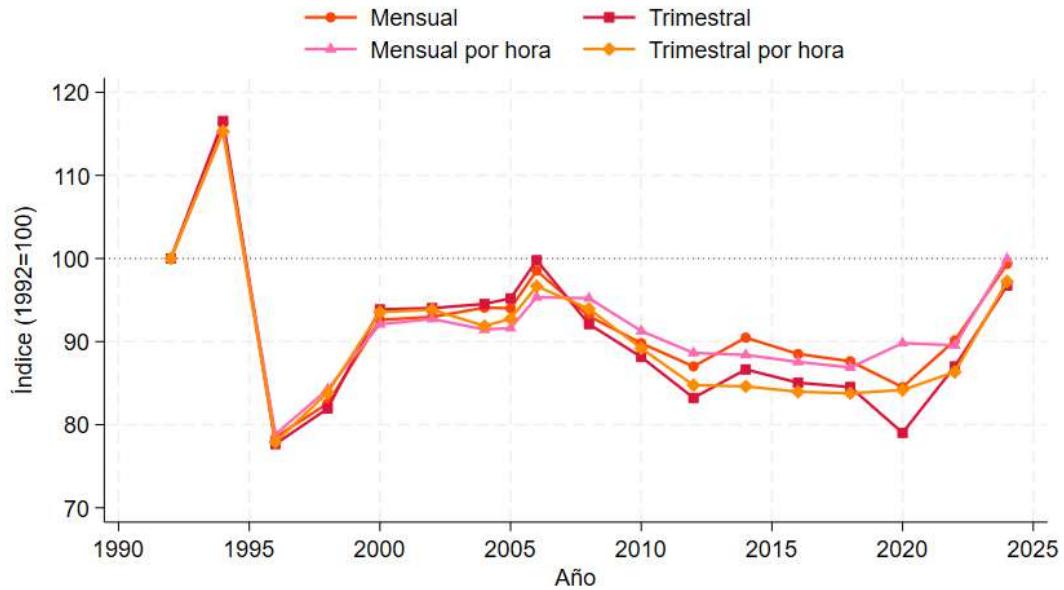
Las gráficas muestran que, a lo largo del periodo 1992–2024, las **brechas salariales** entre **grupos de población** se han mantenido y en algunos casos se han ampliado. Los **grupos con mayor escolaridad** (en especial **hombres y mujeres con preparatoria o más**) presentan de forma consistente los **niveles más altos de ingreso y salario por hora**, mientras que los **grupos con menor escolaridad** concentran los **ingresos más bajos**. A partir de 2018 se observa una **recuperación general** en los salarios, aunque esta ha sido más marcada en los grupos con **mayor nivel educativo**, lo que ha ampliado la **distancia** respecto a los **grupos rezagados**. En consecuencia, la historia que cuentan las gráficas es la **persistencia de desigualdades estructurales** en el mercado laboral mexicano, donde la **educación** aparece como el principal factor que segmenta los **ingresos reales**, reforzado además por **diferencias de género**.

Figura 8: Boxplot, Ingreso Mensual Real por Grupo Poblacional,(ENIGH, 2018–2024)



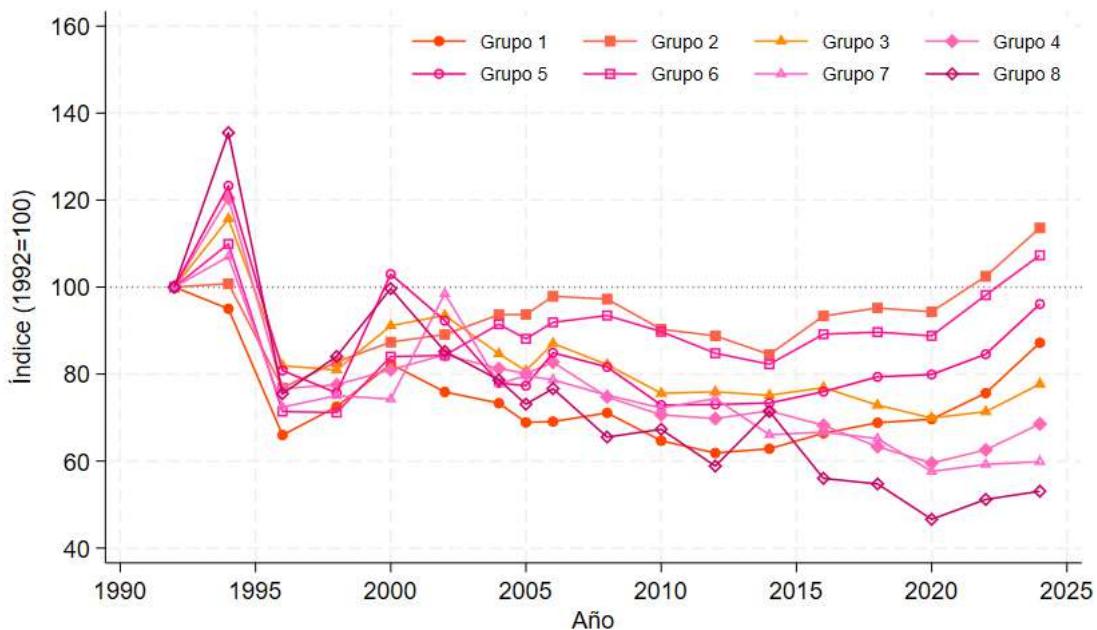
Fuente: Elaboración propia con datos de la ENIGH. Proceso disponible en:  danifuentesga

Figura 9: Tendencias comparadas de salarios reales en México: mensual, trimestral y por hora,(ENIGH, 1992–2024)



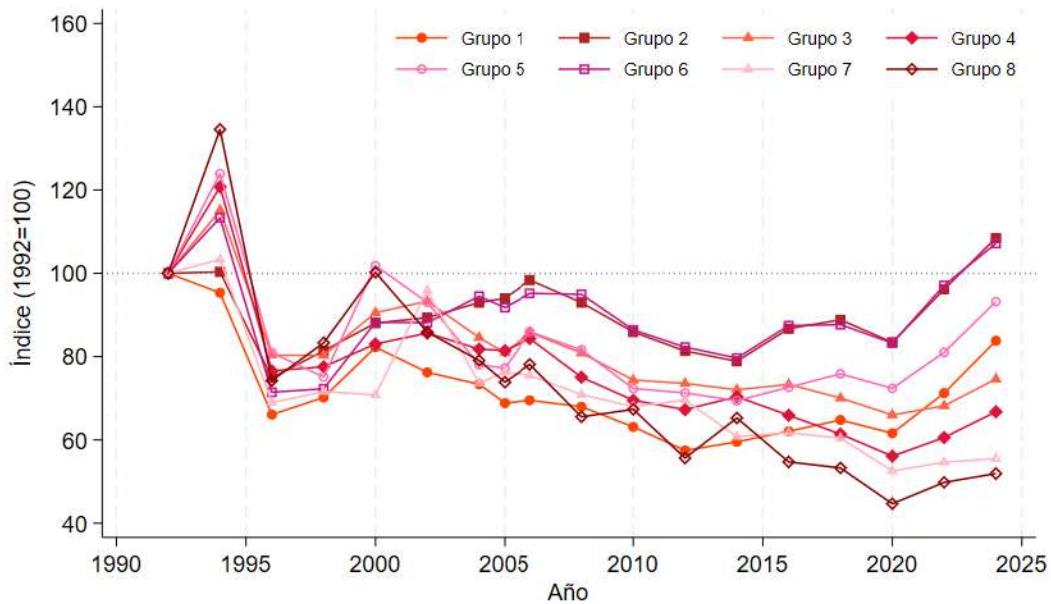
Fuente: Elaboración propia con datos de la ENIGH. Proceso disponible en:  danifuentesga

Figura 10: Evolución del índice del salario mensual real por grupo de población,(ENIGH, 1992–2024)



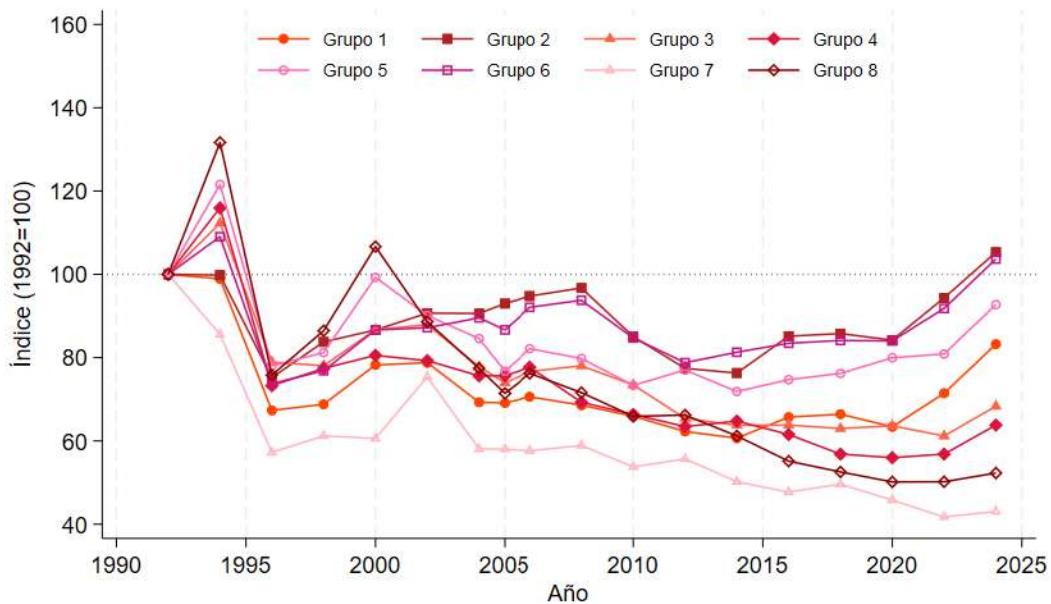
Fuente: Elaboración propia con datos de la ENIGH. Proceso disponible en:  danifuentesga

Figura 11: Evolución del índice del salario trimestral real por grupo de población,(ENIGH, 1992–2024)



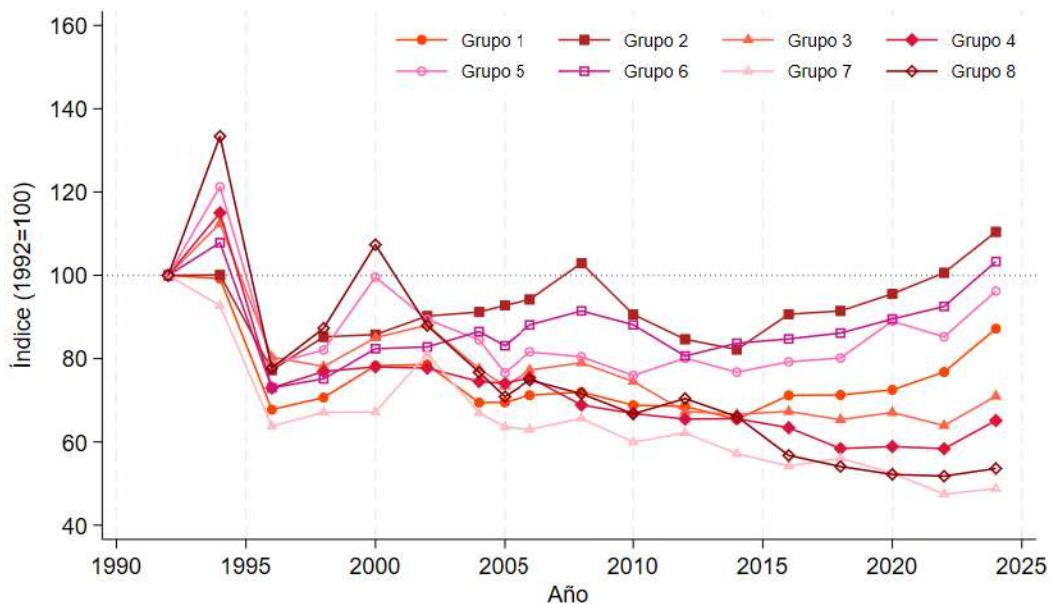
Fuente: Elaboración propia con datos de la ENIGH. Proceso disponible en:  danifuentesga

Figura 12: Evolución del índice del salario trimestral por hora real por grupo de población,(ENIGH, 1992–2024)



Fuente: Elaboración propia con datos de la ENIGH. Proceso disponible en:  danifuentesga

Figura 13: Evolución del índice del salario mensual por hora real por grupo de población,(ENIGH, 1992–2024)



Fuente: Elaboración propia con datos de la ENIGH. Proceso disponible en:  danifuentesga

Durante la **elaboración de las gráficas** me percaté de que los **pesos de ENOE/ENIGH** no son enteros, sino factores de expansión continuos. El comando **fweight** sólo acepta **valores enteros**, por lo que a veces no sirve para el cálculo de **estadísticas descriptivas** como las que usa **CONEVAL o ENOE** en temas de ingreso. En estos casos es más adecuado emplear **aweight o pweight**, que sí permiten trabajar con ponderadores continuos y reflejan mejor la estructura muestral.

**Tabla 7.** Estadísticas descriptivas: mediana de salarios por año y tipo de medición, ENIGH 1992–2024

Año	Salario trimestral	Salario mensual	Salario mensual por hora	Salario trimestral por hora
1992	26968.04	9248.75	48.99	47.63
1994	31430.96	10757.46	56.50	54.93
1996	20946.05	7259.23	38.61	37.11
1998	22096.76	7635.50	41.28	39.94
2000	25314.28	8565.84	45.11	44.56
2002	25359.00	8598.63	45.43	44.70
2004	25491.46	8701.58	44.80	43.76
2005	25676.24	8694.88	44.89	44.18
2006	26911.26	9112.83	46.71	46.04
2008	24839.35	8606.46	46.65	44.75
2010	21305.39	7542.84	40.78	38.15
2012	22445.54	8047.72	43.42	40.37
2014	23368.10	8367.46	43.30	40.30
2016	22937.26	8186.53	42.89	39.99
2018	22795.19	8104.86	42.56	39.91
2020	21311.52	7816.93	44.00	40.10
2022	23464.39	8336.76	43.88	41.13
2024	26089.45	9187.50	48.96	46.32

Fuente: Elaboración propia con datos de la ENIGH. Proceso disponible en:  danifuentesga

**Tabla 8.** Tasas de crecimiento interanual de salarios reales, ENIGH 1992–2024

Año	Salario trimestral	Salario mensual	Salario mensual por hora	Salario trimestral por hora
1994	16.55 %	16.31 %	15.34 %	15.33 %
1996	-33.36 %	-32.52 %	-31.66 %	-32.44 %
1998	5.49 %	5.18 %	6.91 %	7.62 %
2000	14.56 %	12.18 %	9.27 %	11.56 %
2002	0.18 %	0.38 %	0.70 %	0.31 %
2004	0.52 %	1.20 %	-1.38 %	-2.09 %
2005	0.72 %	-0.08 %	0.20 %	0.95 %
2006	4.81 %	4.81 %	4.05 %	4.21 %
2008	-7.70 %	-5.56 %	-0.13 %	-2.81 %
2010	-14.23 %	-12.36 %	-12.59 %	-14.75 %
2012	5.35 %	6.69 %	6.49 %	5.82 %
2014	4.11 %	3.97 %	-0.27 %	-0.17 %
2016	-1.84 %	-2.16 %	-0.95 %	-0.77 %
2018	-0.62 %	-1.00 %	-0.76 %	-0.22 %
2020	-6.51 %	-3.55 %	3.37 %	0.48 %
2022	10.10 %	6.65 %	-0.28 %	2.57 %
2024	11.19 %	10.20 %	11.59 %	12.63 %

Fuente: Elaboración propia con datos de la ENIGH. Proceso disponible en:  danifuentesga

**Tabla 9.** Salario promedio mensual real por grupo de población, ENIGH 1992–2024

Año	Grupo 1	Grupo 2	Grupo 3	Grupo 4	Grupo 5	Grupo 6	Grupo 7	Grupo 8
1992	5927.15	6568.64	12374.18	17405.80	4950.57	6590.33	17397.03	26935.69
1994	5632.02	6620.24	14314.93	20962.58	6103.44	7244.03	18615.98	36489.20
1996	3910.17	5046.99	10137.88	13340.93	4000.82	4704.93	12597.06	20337.92
1998	4297.14	5439.52	10018.02	13499.31	3746.35	4689.00	13053.58	22642.04
2000	4878.72	5737.99	11274.88	14097.08	5098.06	5536.80	12918.37	26840.98
2002	4498.73	5853.05	11569.72	14696.71	4568.29	5556.01	17136.19	22959.11
2004	4345.27	6150.97	10478.11	14149.05	3858.08	6028.11	13531.42	21184.09
2005	4083.79	6153.35	9992.58	13947.82	3828.56	5806.92	13855.72	19686.17
2006	4093.49	6429.51	10770.82	14396.52	4203.23	6056.22	13686.25	20646.12
2008	4213.32	6387.71	10168.89	13010.67	4041.43	6159.94	13074.05	17647.74
2010	3832.99	5932.50	9351.78	12293.46	3607.88	5915.57	12560.35	18130.50
2012	3666.58	5832.23	9394.30	12144.78	3614.49	5588.31	12939.69	15851.46
2014	3724.80	5552.08	9296.90	12452.15	3631.69	5421.24	11493.46	19237.00
2016	3935.55	6132.73	9514.07	11874.87	3760.38	5876.22	11594.93	15091.48
2018	4079.13	6252.02	9015.43	11016.04	3929.13	5908.49	11339.36	14745.85
2020	4126.99	6195.52	8649.65	10364.43	3955.96	5851.16	10023.55	12555.74
2022	4482.19	6731.35	8828.98	10893.01	4186.73	6469.19	10306.32	13788.10
2024	5168.13	7461.26	9614.41	11929.84	4757.31	7071.69	10415.51	14292.89

*Nota:* Los grupos de población se definieron según edad, sexo y nivel educativo: Grupo 1 = Mujeres 25–45 años, < Preparatoria; Grupo 2 = Hombres 25–45 años, < Preparatoria; Grupo 3 = Mujeres 25–45 años, ≥ Preparatoria; Grupo 4 = Hombres 25–45 años, ≥ Preparatoria; Grupo 5 = Mujeres 46–65 años, < Preparatoria; Grupo 6 = Hombres 46–65 años, < Preparatoria; Grupo 7 = Mujeres 46–65 años, ≥ Preparatoria; Grupo 8 = Hombres 46–65 años, ≥ Preparatoria.

Fuente: Elaboración propia con datos de la ENIGH. Proceso disponible en:  danifuentesga

**Tabla 10.** Evolución del salario trimestral real por grupo de población, ENIGH 1992–2024

Año	Grupo 1	Grupo 2	Grupo 3	Grupo 4	Grupo 5	Grupo 6	Grupo 7	Grupo 8
1992	17008.02	19172.91	36560.57	50738.05	14429.73	18672.11	54286.59	79846.95
1994	16214.03	19239.02	42081.93	61285.41	17881.99	21160.75	56105.68	107421.07
1996	11237.37	14361.38	29373.05	38850.69	11677.24	13339.28	37428.71	59288.94
1998	11936.68	15605.77	29360.03	39352.02	10838.72	13494.76	38868.90	66513.96
2000	13990.96	16881.65	33104.49	42125.47	14695.57	16472.48	38430.83	80034.80
2002	12963.46	17134.76	34103.82	43458.40	13426.42	16457.51	52046.37	68573.20
2004	12479.85	17825.29	30924.47	41508.91	11264.43	17645.13	39979.55	63132.92
2005	11708.44	18025.62	29581.39	41326.92	11146.44	17135.17	41409.20	58963.89
2006	11824.04	18865.25	31376.68	42793.15	12405.08	17774.31	40985.10	62388.05
2008	11561.53	17826.27	29570.46	38080.32	11779.07	17732.35	38489.22	52339.64
2010	10735.42	16472.32	27188.78	35272.92	10437.96	16136.03	36893.06	53770.70
2012	9775.87	15597.85	26892.79	34131.91	10285.21	15362.81	37763.36	44424.12
2014	10123.51	15122.95	26338.18	35702.74	10015.43	14872.96	32998.81	52111.30
2016	10555.31	16624.63	26807.15	33444.26	10479.65	16328.60	33505.33	43693.08
2018	11017.05	17033.58	25606.78	31150.42	10945.15	16369.82	32809.06	42529.23
2020	10476.52	15995.33	24124.14	28468.03	10443.20	15547.49	28527.85	35672.36
2022	12114.03	18431.07	24930.99	30733.75	11692.68	18131.89	29666.02	39771.38
2024	14255.88	20798.98	27278.12	33870.68	13447.50	20010.61	30123.87	41456.99

*Nota:* Los grupos de población se definieron según edad, sexo y nivel educativo: Grupo 1 = Mujeres 25–45 años, < Preparatoria; Grupo 2 = Hombres 25–45 años, < Preparatoria; Grupo 3 = Mujeres 25–45 años, ≥ Preparatoria; Grupo 4 = Hombres 25–45 años, ≥ Preparatoria; Grupo 5 = Mujeres 46–65 años, < Preparatoria; Grupo 6 = Hombres 46–65 años, < Preparatoria; Grupo 7 = Mujeres 46–65 años, ≥ Preparatoria; Grupo 8 = Hombres 46–65 años, ≥ Preparatoria.

Fuente: Elaboración propia con datos de la ENIGH. Proceso disponible en:  danifuentesga

**Tabla 11.** Salario mensual por hora promedio según grupo de población, ENIGH 1992–2024

Año	Grupo 1	Grupo 2	Grupo 3	Grupo 4	Grupo 5	Grupo 6	Grupo 7	Grupo 8
1992	34.76	31.06	81.03	92.47	31.40	33.27	128.84	136.17
1994	34.49	31.08	91.13	106.35	38.08	35.88	119.46	181.60
1996	23.55	23.99	65.12	67.48	24.71	24.28	82.17	105.85
1998	24.56	26.47	63.30	71.11	25.77	25.00	86.44	118.87
2000	27.22	26.64	68.88	72.16	31.27	27.40	86.58	146.17
2002	27.31	28.02	71.33	71.86	28.07	27.55	104.53	119.74
2004	24.14	28.32	62.84	68.89	26.53	28.78	86.32	104.47
2005	24.16	28.82	59.53	68.47	24.05	27.64	81.95	96.50
2006	24.75	29.26	62.58	69.80	25.63	29.31	81.12	101.90
2008	25.02	31.97	64.01	63.68	25.26	30.43	84.58	97.45
2010	23.92	28.14	60.37	61.79	23.85	29.32	77.24	90.85
2012	23.82	26.29	54.53	60.55	25.15	26.82	80.12	95.88
2014	22.75	25.53	53.99	60.63	24.09	27.84	73.70	90.00
2016	24.74	28.16	54.55	58.64	24.88	28.19	69.83	77.29
2018	24.78	28.40	52.95	54.01	25.17	28.66	72.29	73.65
2020	25.20	29.68	54.34	54.45	27.94	29.78	67.68	71.08
2022	26.69	31.25	51.80	53.98	26.76	30.78	61.14	70.51
2024	30.30	34.30	57.51	60.23	30.21	34.37	62.85	73.03

*Nota:* Los grupos de población se definieron según edad, sexo y nivel educativo: Grupo 1 = Mujeres 25–45 años, < Preparatoria; Grupo 2 = Hombres 25–45 años, < Preparatoria; Grupo 3 = Mujeres 25–45 años, ≥ Preparatoria; Grupo 4 = Hombres 25–45 años, ≥ Preparatoria; Grupo 5 = Mujeres 46–65 años, < Preparatoria; Grupo 6 = Hombres 46–65 años, < Preparatoria; Grupo 7 = Mujeres 46–65 años, ≥ Preparatoria; Grupo 8 = Hombres 46–65 años, ≥ Preparatoria.

Fuente: Elaboración propia con datos de la ENIGH. Proceso disponible en:  danifuentesga

**Tabla 12.** Salario trimestral por hora promedio según grupo de población, ENIGH 1992–2024

Año	Grupo 1	Grupo 2	Grupo 3	Grupo 4	Grupo 5	Grupo 6	Grupo 7	Grupo 8
1992	33.30	30.28	79.50	89.40	30.57	31.30	140.39	134.93
1994	32.94	30.22	89.26	103.65	37.16	34.13	120.17	177.66
1996	22.42	22.77	62.76	65.50	23.85	23.08	80.35	102.29
1998	22.91	25.36	62.05	69.22	24.82	24.06	85.95	116.66
2000	26.05	26.23	68.94	72.01	30.34	27.12	85.07	143.90
2002	26.24	27.46	69.87	70.87	27.55	27.29	105.80	119.56
2004	23.07	27.43	61.86	67.60	25.85	28.03	81.57	104.48
2005	23.01	28.15	58.80	67.63	23.46	27.12	81.45	96.36
2006	23.51	28.70	60.95	69.46	25.11	28.82	80.92	102.84
2008	22.83	29.28	62.04	61.94	24.39	29.35	82.70	96.61
2010	21.96	25.75	58.38	59.32	22.42	26.52	75.46	88.92
2012	20.74	23.44	52.01	56.73	23.54	24.66	78.18	89.34
2014	20.20	23.10	50.67	57.91	21.96	25.45	70.50	82.57
2016	21.88	25.78	50.74	55.02	22.84	26.12	66.98	74.40
2018	22.12	25.97	50.07	50.81	23.30	26.33	69.67	70.91
2020	21.12	25.49	50.59	50.04	24.44	26.30	64.28	67.67
2022	23.79	28.57	48.65	50.81	24.72	28.74	58.61	67.73
2024	27.72	31.89	54.32	57.04	28.34	32.46	60.43	70.55

*Nota:* Los grupos de población se definieron según edad, sexo y nivel educativo: Grupo 1 = Mujeres 25–45 años, < Preparatoria; Grupo 2 = Hombres 25–45 años, < Preparatoria; Grupo 3 = Mujeres 25–45 años, ≥ Preparatoria; Grupo 4 = Hombres 25–45 años, ≥ Preparatoria; Grupo 5 = Mujeres 46–65 años, < Preparatoria; Grupo 6 = Hombres 46–65 años, < Preparatoria; Grupo 7 = Mujeres 46–65 años, ≥ Preparatoria; Grupo 8 = Hombres 46–65 años, ≥ Preparatoria.

Fuente: Elaboración propia con datos de la ENIGH. Proceso disponible en:  danifuentesga

**Tabla 13.** Crecimiento porcentual del salario mensual real promedio según grupo de población, ENIGH 1992–2024

Año	Grupo 1	Grupo 2	Grupo 3	Grupo 4	Grupo 5	Grupo 6	Grupo 7	Grupo 8
1994	-4.98	0.79	15.68	20.43	23.29	9.92	7.01	35.47
1996	-30.57	-23.76	-29.18	-36.36	-34.45	-35.05	-32.33	-44.26
1998	9.90	7.78	-1.18	1.19	-6.36	-0.34	3.62	11.33
2000	13.53	5.49	12.55	4.43	36.08	18.08	-1.04	18.54
2002	-7.79	2.01	2.61	4.25	-10.39	0.35	32.65	-14.46
2004	-3.41	5.09	-9.44	-3.73	-15.55	8.50	-21.04	-7.73
2005	-6.02	0.04	-4.63	-1.42	-0.77	-3.67	2.40	-7.07
2006	0.24	4.49	7.79	3.22	9.79	4.29	-1.22	4.88
2008	2.93	-0.65	-5.59	-9.63	-3.85	1.71	-4.47	-14.52
2010	-9.03	-7.13	-8.04	-5.51	-10.73	-3.97	-3.93	2.74
2012	-4.34	-1.69	0.45	-1.21	0.18	-5.53	3.02	-12.57
2014	1.59	-4.80	-1.04	2.53	0.48	-2.99	-11.18	21.36
2016	5.66	10.46	2.34	-4.64	3.54	8.39	0.88	-21.55
2018	3.65	1.95	-5.24	-7.23	4.49	0.55	-2.20	-2.29
2020	1.17	-0.90	-4.06	-5.92	0.68	-0.97	-11.60	-14.85
2022	8.61	8.65	2.07	5.10	5.83	10.56	2.82	9.82
2024	15.30	10.84	8.90	9.52	13.63	9.31	1.06	3.66

*Nota:* Los grupos de población se definieron según edad, sexo y nivel educativo: Grupo 1 = Mujeres 25–45 años, < Preparatoria; Grupo 2 = Hombres 25–45 años, < Preparatoria; Grupo 3 = Mujeres 25–45 años, ≥ Preparatoria; Grupo 4 = Hombres 25–45 años, ≥ Preparatoria; Grupo 5 = Mujeres 46–65 años, < Preparatoria; Grupo 6 = Hombres 46–65 años, < Preparatoria; Grupo 7 = Mujeres 46–65 años, ≥ Preparatoria; Grupo 8 = Hombres 46–65 años, ≥ Preparatoria.

Fuente: Elaboración propia con datos de la ENIGH. Proceso disponible en:  danifuentesga

**Tabla 14.** Crecimiento porcentual del salario trimestral real promedio según grupo de población, ENIGH 1992–2024

Año	Grupo 1	Grupo 2	Grupo 3	Grupo 4	Grupo 5	Grupo 6	Grupo 7	Grupo 8
1994	-4.67	0.34	15.10	20.79	23.92	13.33	3.35	34.53
1996	-30.69	-25.35	-30.20	-36.61	-34.70	-36.96	-33.29	-44.81
1998	6.22	8.66	-0.04	1.29	-7.18	1.17	3.85	12.19
2000	17.21	8.18	12.75	7.05	35.58	22.07	-1.13	20.33
2002	-7.34	1.50	3.02	3.16	-8.64	-0.09	35.43	-14.32
2004	-3.73	4.03	-9.32	-4.49	-16.10	7.22	-23.18	-7.93
2005	-6.18	1.12	-4.34	-0.44	-1.05	-2.89	3.58	-6.60
2006	0.99	4.66	6.07	3.55	11.29	3.73	-1.02	5.81
2008	-2.22	-5.51	-5.76	-11.01	-5.05	-0.24	-6.09	-16.11
2010	-7.15	-7.60	-8.05	-7.37	-11.39	-9.00	-4.15	2.73
2012	-8.94	-5.31	-1.09	-3.23	-1.46	-4.79	2.36	-17.38
2014	3.56	-3.04	-2.06	4.60	-2.62	-3.19	-12.62	17.30
2016	4.27	9.93	1.78	-6.33	4.64	9.79	1.53	-16.15
2018	4.37	2.46	-4.48	-6.86	4.44	0.25	-2.08	-2.66
2020	-4.91	-6.10	-5.79	-8.61	-4.59	-5.02	-13.05	-16.12
2022	15.63	15.23	3.34	7.96	11.96	16.62	3.99	11.49
2024	17.68	12.85	9.41	10.21	15.01	10.36	1.54	4.24

*Nota:* Los grupos de población se definieron según edad, sexo y nivel educativo: Grupo 1 = Mujeres 25–45 años, < Preparatoria; Grupo 2 = Hombres 25–45 años, < Preparatoria; Grupo 3 = Mujeres 25–45 años, ≥ Preparatoria; Grupo 4 = Hombres 25–45 años, ≥ Preparatoria; Grupo 5 = Mujeres 46–65 años, < Preparatoria; Grupo 6 = Hombres 46–65 años, < Preparatoria; Grupo 7 = Mujeres 46–65 años, ≥ Preparatoria; Grupo 8 = Hombres 46–65 años, ≥ Preparatoria.

Fuente: Elaboración propia con datos de la ENIGH. Proceso disponible en:  danifuentesga

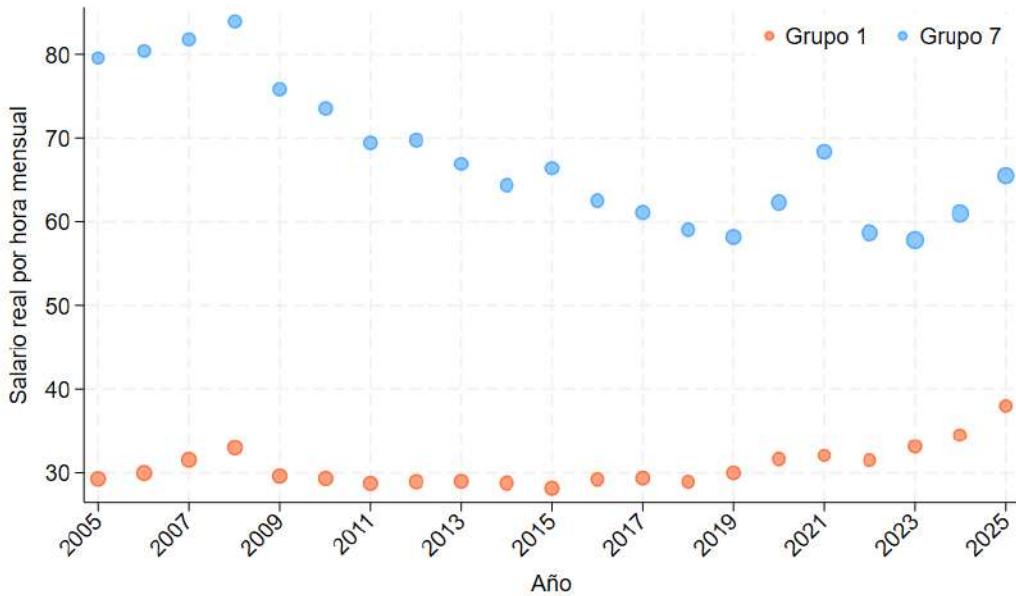
**Tabla 15.** Crecimiento porcentual del salario trimestral por hora real promedio según grupo de población, ENIGH 1992–2024

Año	Grupo 1	Grupo 2	Grupo 3	Grupo 4	Grupo 5	Grupo 6	Grupo 7	Grupo 8
1994	-0.79	0.07	12.46	15.01	21.27	7.84	-7.28	33.37
1996	-31.72	-22.81	-28.54	-36.55	-35.12	-32.32	-31.22	-41.71
1998	4.31	10.32	-2.80	5.38	4.31	2.97	5.20	12.30
2000	10.81	0.66	8.82	1.47	21.32	9.60	0.16	22.96
2002	0.35	5.18	3.55	-0.41	-10.23	0.54	20.74	-18.08
2004	-11.62	1.07	-11.90	-4.14	-5.47	4.45	-17.42	-12.76
2005	0.10	1.74	-5.27	-0.61	-9.35	-3.95	-5.07	-7.63
2006	2.44	1.53	5.14	1.93	6.54	6.05	-1.02	5.60
2008	1.08	9.27	2.28	-8.77	-1.44	3.81	4.27	-4.36
2010	-4.40	-11.97	-5.68	-2.96	-5.55	-3.66	-8.68	-6.78
2012	-0.41	-6.59	-9.68	-2.01	5.44	-8.51	3.72	5.54
2014	-4.51	-2.91	-0.99	0.14	-4.20	3.80	-8.00	-6.14
2016	8.75	10.32	1.04	-3.29	3.26	1.24	-5.25	-14.12
2018	0.16	0.86	-2.95	-7.89	1.16	1.66	3.51	-4.70
2020	1.71	4.49	2.64	0.82	11.02	3.92	-6.37	-3.50
2022	5.89	5.29	-4.68	-0.88	-4.21	3.37	-9.67	-0.80
2024	13.56	9.78	11.03	11.59	12.89	11.66	2.79	3.57

*Nota:* Los grupos de población se definieron según edad, sexo y nivel educativo: Grupo 1 = Mujeres 25–45 años, < Preparatoria; Grupo 2 = Hombres 25–45 años, < Preparatoria; Grupo 3 = Mujeres 25–45 años, ≥ Preparatoria; Grupo 4 = Hombres 25–45 años, ≥ Preparatoria; Grupo 5 = Mujeres 46–65 años, < Preparatoria; Grupo 6 = Hombres 46–65 años, < Preparatoria; Grupo 7 = Mujeres 46–65 años, ≥ Preparatoria; Grupo 8 = Hombres 46–65 años, ≥ Preparatoria.

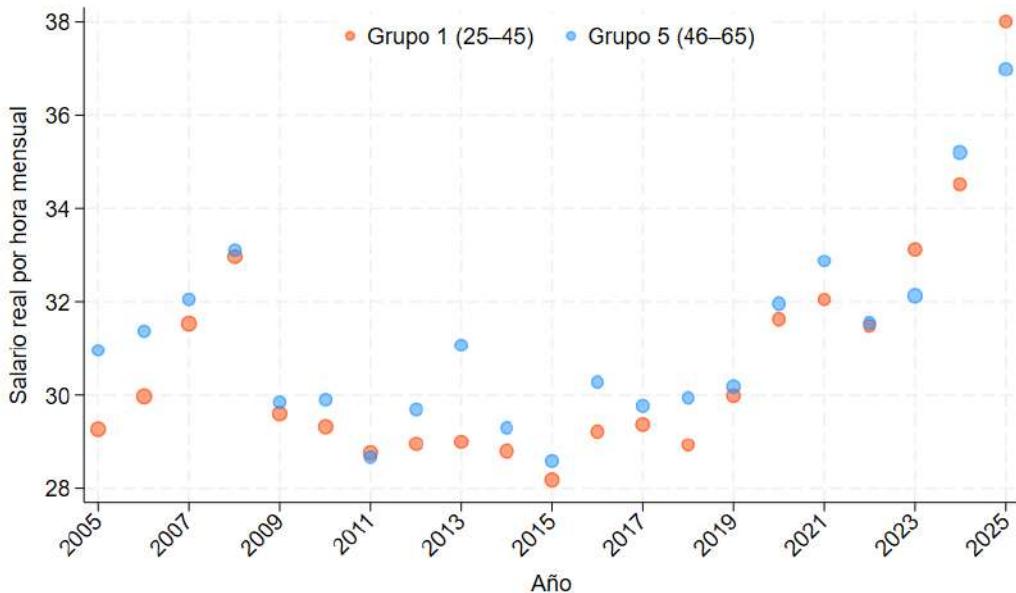
Fuente: Elaboración propia con datos de la ENIGH. Proceso disponible en:  danifuentesga

Figura 14: Evolución del salario real por hora mensual en el primer trimestre: comparación entre Grupo 1 y Grupo 7 (ENOE, Q1, 2005–2025)



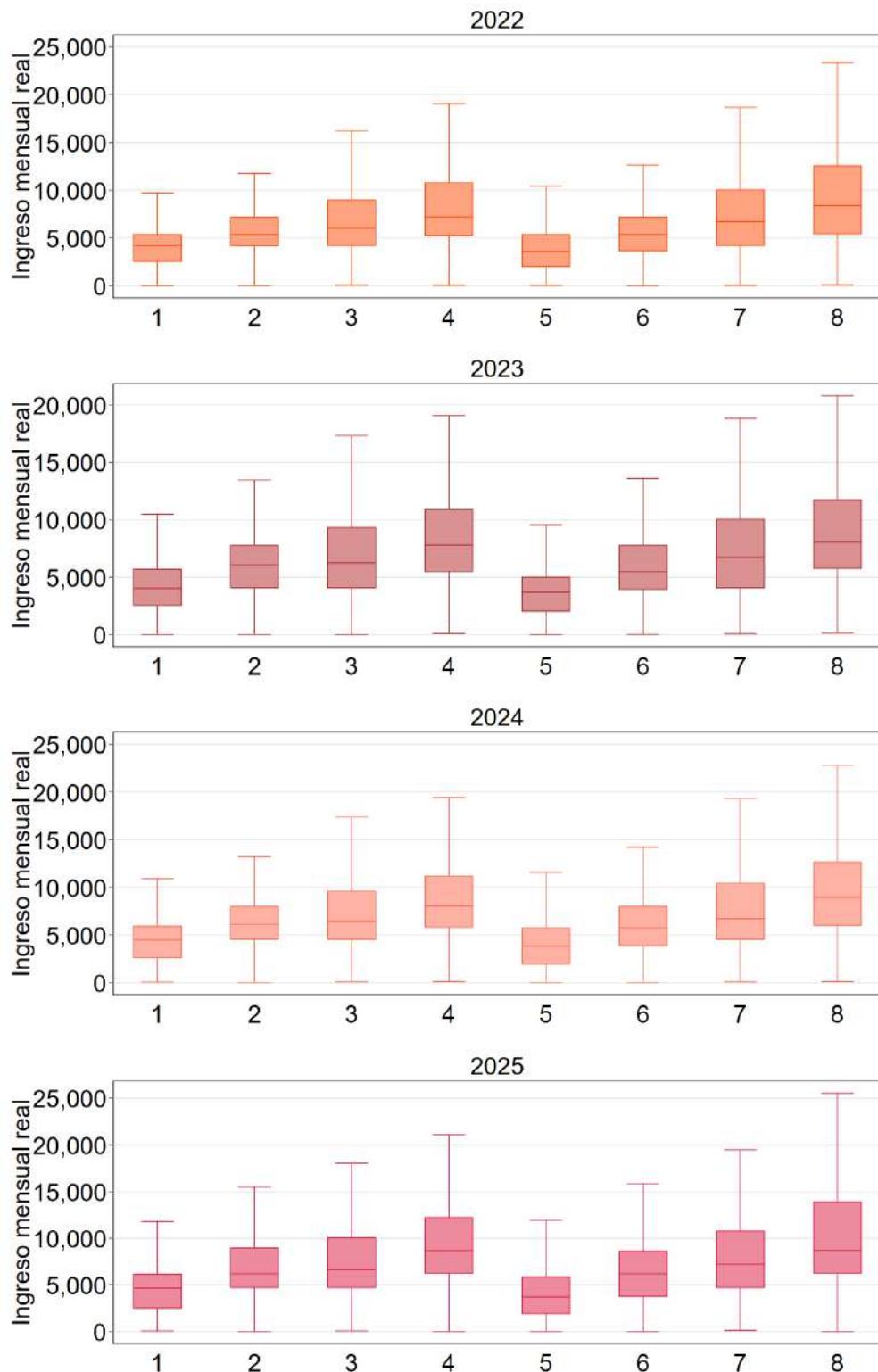
Fuente: Elaboración propia con datos de la ENOE. Proceso disponible en:  danifuentesga

Figura 15: Evolución del salario real por hora mensual en Mujeres de baja educación: comparación entre 25–45 y 46–65 años (ENOE, Q1, 2005–2025)



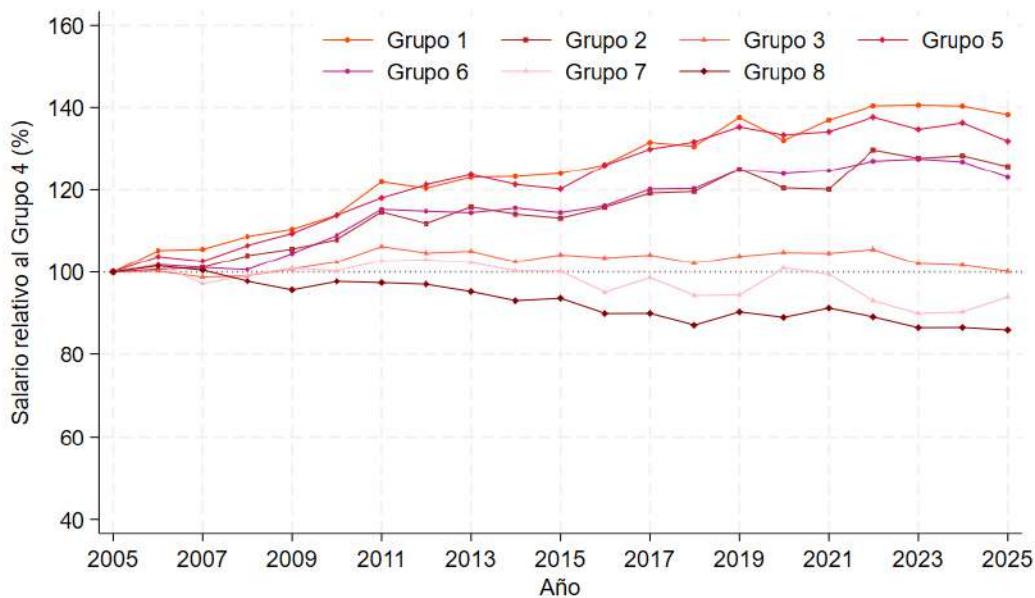
Fuente: Elaboración propia con datos de la ENOE. Proceso disponible en:  danifuentesga

Figura 16: Boxplot, Ingreso Mensual Real por Grupo Poblacional,(ENOE, Q1, 2022-2025)



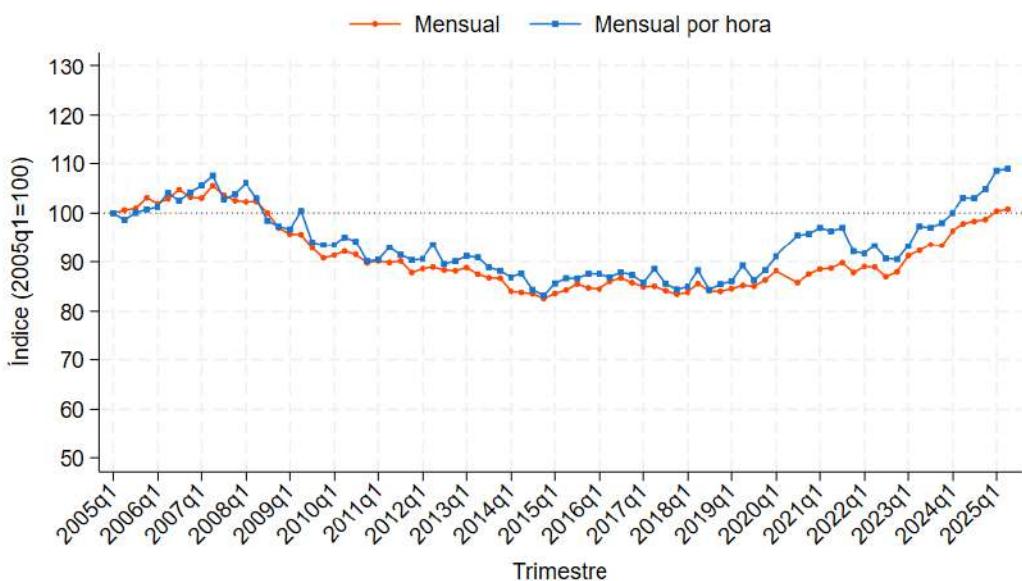
Fuente: Elaboración propia con datos de la ENOE. Proceso disponible en:  danifuentesga

Figura 17: Brechas relativas en salario real por hora entre Grupos; Base Grupo 4=100,(ENOE, 2005–2025)



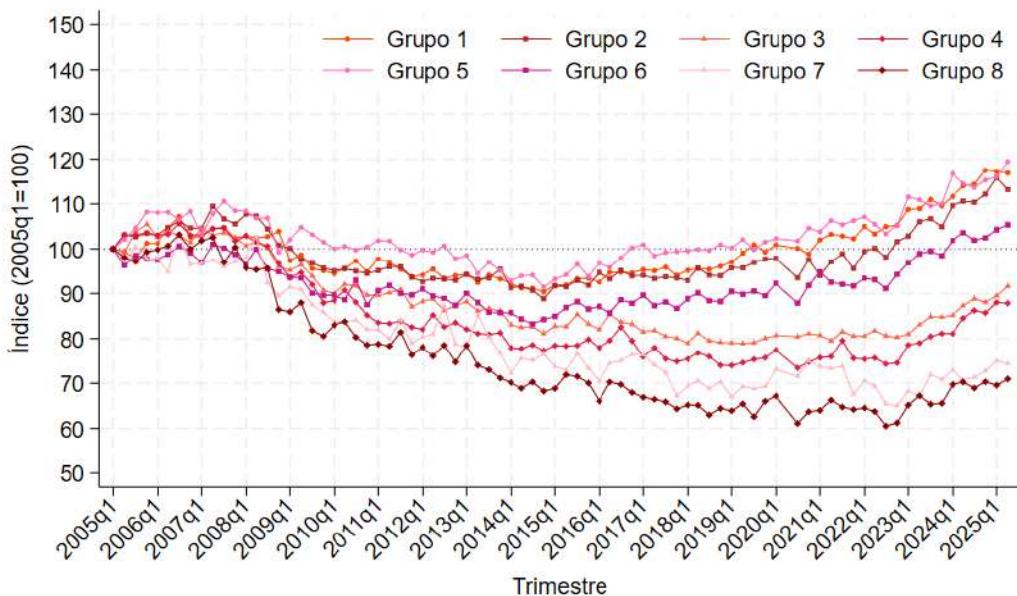
Fuente: Elaboración propia con datos de la ENOE. Proceso disponible en:  danifuentesga

Figura 18: Índices de salarios reales en la ENOE; 2005q1=100,(ENOE, 2005–2025)



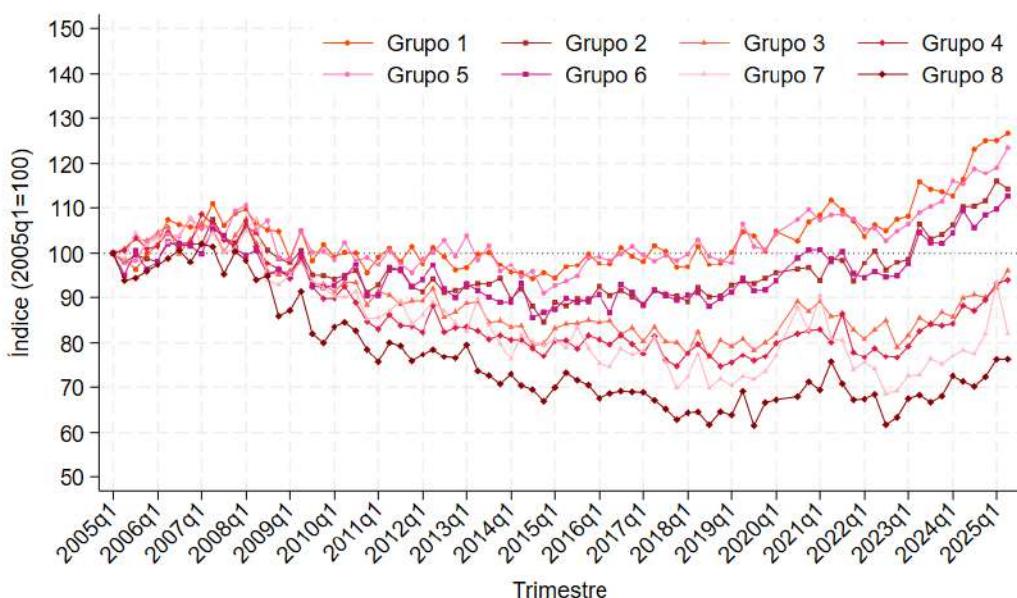
Fuente: Elaboración propia con datos de la ENOE. Proceso disponible en:  danifuentesga

Figura 19: Índices de salario mensual por grupo de población en la ENOE; 2005q1=100,(ENOE, 2005–2025)



Fuente: Elaboración propia con datos de la ENOE. Proceso disponible en:  danifuentesga

Figura 20: Índices de salario por hora por grupo de población en la ENOE; 2005q1=100,(ENOE, 2005–2025)



Fuente: Elaboración propia con datos de la ENOE. Proceso disponible en:  danifuentesga

**Tabla 16.** Salario mensual y por hora promedio en la ENOE (2005–2025)

Año	Salario mensual	Salario por hora
2005	7786.64	45.47
2006	7939.36	46.90
2007	7977.59	47.78
2008	7727.05	46.08
2009	7213.93	43.77
2010	7015.54	42.46
2011	6885.35	41.56
2012	6809.96	41.41
2013	6726.81	40.87
2014	6422.74	38.92
2015	6502.67	39.43
2016	6597.35	39.78
2017	6493.40	39.19
2018	6491.77	39.04
2019	6562.57	39.83
2020	6710.09	42.80
2021	6825.14	43.52
2022	6788.77	41.68
2023	7125.52	43.88
2024	7525.16	46.76
2025	7740.21	49.51

Fuente: Elaboración propia con datos de la ENOE. Proceso disponible en:  danifuentesga

**Tabla 17.** Crecimiento interanual del salario mensual y por hora en la ENOE (2005–2025)

Año	Crec. salario mensual (%)	Crec. salario por hora (%)
2006	1.96	3.15
2007	0.48	1.87
2008	-3.14	-3.54
2009	-6.64	-5.01
2010	-2.75	-3.01
2011	-1.86	-2.11
2012	-1.10	-0.37
2013	-1.22	-1.29
2014	-4.52	-4.77
2015	1.24	1.31
2016	1.46	0.89
2017	-1.58	-1.51
2018	-0.03	-0.38
2019	1.09	2.02
2020	2.25	7.48
2021	1.71	1.68
2022	-0.53	-4.24
2023	4.96	5.29
2024	5.61	6.56
2025	2.86	5.88

Fuente: Elaboración propia con datos de la ENOE. Proceso disponible en:  danifuentesga

**Tabla 18.** Salario mensual promedio por Grupos en la ENOE (2005–2025)

Año	Grupo 1	Grupo 2	Grupo 3	Grupo 4	Grupo 5	Grupo 6	Grupo 7	Grupo 8
2005	4059.17	6358.19	9350.37	12388.51	3864.66	6207.58	11558.88	16098.18
2006	4233.66	6502.02	9397.60	12541.02	4015.01	6263.36	11458.95	16468.19
2007	4219.94	6622.16	9462.88	12514.35	4006.00	6276.93	11310.15	16391.44
2008	4196.25	6530.45	9113.38	12152.61	3925.88	6116.81	11165.74	15221.48
2009	3937.18	6061.91	8604.49	11130.48	3831.90	5801.07	10386.70	13708.69
2010	3891.74	5904.50	8312.76	10654.92	3732.18	5674.20	9738.12	13291.59
2011	3908.90	5913.45	8177.77	10077.36	3743.10	5728.81	9488.14	12844.50
2012	3839.05	5781.02	8021.87	10078.10	3701.57	5643.71	9542.23	12535.62
2013	3810.78	5844.96	7939.41	9837.09	3586.58	5530.81	9351.81	12108.52
2014	3718.41	5629.46	7531.86	9418.75	3466.35	5339.23	8738.55	11322.02
2015	3773.08	5716.37	7638.60	9525.94	3517.75	5480.76	8661.57	11517.53
2016	3835.17	5856.47	7648.92	9660.05	3640.87	5523.48	8654.40	11180.19
2017	3873.23	5819.28	7408.90	9194.78	3706.76	5564.15	8475.98	10751.00
2018	3892.64	5850.36	7293.37	9132.19	3724.70	5623.82	8149.06	10513.66
2019	4036.22	5997.40	7255.40	9055.72	3758.63	5699.40	8016.70	10528.40
2020	4073.13	5990.04	7388.88	9098.75	3833.94	5741.85	8557.25	10461.22
2021	4179.62	5987.81	7374.11	9269.48	3928.87	5868.22	8405.94	10574.14
2022	4262.01	6201.92	7393.45	9063.71	3922.10	5881.32	7890.46	10201.02
2023	4463.54	6531.79	7633.43	9649.77	4116.17	6227.72	8126.15	10741.50
2024	4658.68	6875.63	7991.47	10202.49	4285.55	6482.43	8400.02	11392.06
2025	4768.99	7114.27	8286.16	10633.34	4384.03	6630.29	8724.61	11463.39

*Nota:* Los grupos de población se definieron según edad, sexo y nivel educativo: Grupo 1 = Mujeres 25–45 años, < Preparatoria; Grupo 2 = Hombres 25–45 años, < Preparatoria; Grupo 3 = Mujeres 25–45 años, ≥ Preparatoria; Grupo 4 = Hombres 25–45 años, ≥ Preparatoria; Grupo 5 = Mujeres 46–65 años, < Preparatoria; Grupo 6 = Hombres 46–65 años, < Preparatoria; Grupo 7 = Mujeres 46–65 años, ≥ Preparatoria; Grupo 8 = Hombres 46–65 años, ≥ Preparatoria.

Fuente: Elaboración propia con datos de la ENOE. Proceso disponible en:  danifuentesga

**Tabla 19.** Salario por hora promedio por Grupo en la ENOE (2005–2025)

Año	Grupo 1	Grupo 2	Grupo 3	Grupo 4	Grupo 5	Grupo 6	Grupo 7	Grupo 8
2005	28.36	33.42	63.33	66.13	28.97	34.36	79.96	90.37
2006	30.28	34.09	64.34	67.02	30.50	35.45	82.61	92.98
2007	31.03	35.00	64.78	68.50	30.82	35.98	80.48	94.01
2008	30.63	34.57	62.34	65.71	30.66	34.32	78.78	87.81
2009	29.02	32.73	59.11	61.28	29.38	33.31	74.76	80.10
2010	28.38	31.69	57.00	58.08	28.95	32.90	70.43	77.56
2011	28.74	31.85	55.96	54.97	28.41	32.92	68.30	73.18
2012	28.32	31.02	55.06	54.89	29.15	32.75	68.44	72.79
2013	28.30	31.41	54.03	53.65	29.06	31.90	66.37	69.85
2014	27.36	29.86	50.85	51.77	27.50	31.10	62.77	65.83
2015	27.89	30.00	52.40	52.47	27.62	31.21	63.57	67.15
2016	28.42	30.74	52.08	52.53	29.00	31.70	60.39	64.59
2017	28.52	30.45	50.45	50.55	28.77	31.56	60.29	62.19
2018	28.26	30.49	49.38	50.45	29.10	31.53	57.53	60.10
2019	29.41	31.50	49.60	49.84	29.54	32.34	56.93	61.55
2020	30.15	32.41	53.46	53.24	31.10	34.29	65.01	64.78
2021	31.38	32.42	53.54	53.41	31.38	34.61	64.20	66.63
2022	30.35	33.07	50.98	50.42	30.41	33.29	56.72	61.44
2023	32.44	34.76	52.60	53.83	31.77	35.68	58.58	63.70
2024	34.26	36.95	55.44	56.94	33.99	37.54	62.20	67.40
2025	36.18	38.78	58.53	61.02	35.25	39.00	69.35	71.74

*Nota:* Los grupos de población se definieron según edad, sexo y nivel educativo: Grupo 1 = Mujeres 25–45 años, < Preparatoria; Grupo 2 = Hombres 25–45 años, < Preparatoria; Grupo 3 = Mujeres 25–45 años, ≥ Preparatoria; Grupo 4 = Hombres 25–45 años, ≥ Preparatoria; Grupo 5 = Mujeres 46–65 años, < Preparatoria; Grupo 6 = Hombres 46–65 años, < Preparatoria; Grupo 7 = Mujeres 46–65 años, ≥ Preparatoria; Grupo 8 = Hombres 46–65 años, ≥ Preparatoria.

Fuente: Elaboración propia con datos de la ENOE. Proceso disponible en:  danifuentesga

**Tabla 20.** Crecimiento interanual del salario mensual por Grupo en la ENOE (2005–2025)

Año	Grupo 1	Grupo 2	Grupo 3	Grupo 4	Grupo 5	Grupo 6	Grupo 7	Grupo 8
2006	4.30	2.26	0.51	1.23	3.89	0.90	-0.86	2.30
2007	-0.32	1.85	0.69	-0.21	-0.22	0.22	-1.30	-0.47
2008	-0.56	-1.38	-3.69	-2.89	-2.00	-2.55	-1.28	-7.14
2009	-6.17	-7.17	-5.58	-8.41	-2.39	-5.16	-6.98	-9.94
2010	-1.15	-2.60	-3.39	-4.27	-2.60	-2.19	-6.24	-3.04
2011	0.44	0.15	-1.62	-5.42	0.29	0.96	-2.57	-3.36
2012	-1.79	-2.24	-1.91	0.01	-1.11	-1.49	0.57	-2.40
2013	-0.74	1.11	-1.03	-2.39	-3.11	-2.00	-2.00	-3.41
2014	-2.42	-3.69	-5.13	-4.25	-3.35	-3.46	-6.56	-6.50
2015	1.47	1.54	1.42	1.14	1.48	2.65	-0.88	1.73
2016	1.65	2.45	0.14	1.41	3.50	0.78	-0.08	-2.93
2017	0.99	-0.63	-3.14	-4.82	1.81	0.74	-2.06	-3.84
2018	0.50	0.53	-1.56	-0.68	0.48	1.07	-3.86	-2.21
2019	3.69	2.51	-0.52	-0.84	0.91	1.34	-1.62	0.14
2020	0.91	-0.12	1.84	0.48	2.00	0.74	6.74	-0.64
2021	2.61	-0.04	-0.20	1.88	2.48	2.20	-1.77	1.08
2022	1.97	3.58	0.26	-2.22	-0.17	0.22	-6.13	-3.53
2023	4.73	5.32	3.25	6.47	4.95	5.89	2.99	5.30
2024	4.37	5.26	4.69	5.73	4.12	4.09	3.37	6.06
2025	2.37	3.47	3.69	4.22	2.30	2.28	3.86	0.63

*Nota:* Los grupos de población se definieron según edad, sexo y nivel educativo: Grupo 1 = Mujeres 25–45 años, < Preparatoria; Grupo 2 = Hombres 25–45 años, < Preparatoria; Grupo 3 = Mujeres 25–45 años,  $\geq$  Preparatoria; Grupo 4 = Hombres 25–45 años,  $\geq$  Preparatoria; Grupo 5 = Mujeres 46–65 años, < Preparatoria; Grupo 6 = Hombres 46–65 años, < Preparatoria; Grupo 7 = Mujeres 46–65 años,  $\geq$  Preparatoria; Grupo 8 = Hombres 46–65 años,  $\geq$  Preparatoria.

Fuente: Elaboración propia con datos de la ENOE. Proceso disponible en:  danifuentesga

**Tabla 21.** Crecimiento porcentual del salario mensual por hora según grupo poblacional, ENOE (2005–2025)

Año	Grupo 1	Grupo 2	Grupo 3	Grupo 4	Grupo 5	Grupo 6	Grupo 7	Grupo 8
2006	6.76	2.00	1.59	1.35	5.29	3.18	3.31	2.89
2007	2.48	2.66	0.68	2.21	1.06	1.49	-2.57	1.11
2008	-1.27	-1.22	-3.76	-4.07	-0.52	-4.60	-2.11	-6.60
2009	-5.26	-5.33	-5.18	-6.75	-4.19	-2.96	-5.11	-8.77
2010	-2.22	-3.18	-3.58	-5.22	-1.44	-1.23	-5.79	-3.17
2011	1.27	0.50	-1.82	-5.36	-1.87	0.09	-3.02	-5.66
2012	-1.45	-2.58	-1.62	-0.14	2.59	-0.54	0.20	-0.53
2013	-0.05	1.25	-1.87	-2.26	-0.28	-2.57	-3.02	-4.04
2014	-3.35	-4.96	-5.88	-3.50	-5.37	-2.54	-5.43	-5.75
2015	1.96	0.48	3.05	1.35	0.43	0.38	1.27	1.99
2016	1.90	2.46	-0.62	0.11	4.98	1.56	-5.01	-3.81
2017	0.35	-0.94	-3.12	-3.77	-0.80	-0.46	-0.16	-3.72
2018	-0.94	0.16	-2.13	-0.19	1.16	-0.07	-4.58	-3.35
2019	4.08	3.29	0.46	-1.21	1.50	2.55	-1.05	2.41
2020	2.51	2.88	7.77	6.82	5.29	6.03	14.20	5.25
2021	4.07	0.04	0.14	0.32	0.91	0.93	-1.24	2.86
2022	-3.26	2.00	-4.77	-5.60	-3.11	-3.81	-11.65	-7.80
2023	6.89	5.11	3.18	6.77	4.47	7.18	3.27	3.69
2024	5.61	6.29	5.40	5.78	7.01	5.23	6.19	5.82
2025	5.61	4.98	5.56	7.17	3.69	3.89	11.49	6.43

*Nota:* Los grupos de población se definieron según edad, sexo y nivel educativo: Grupo 1 = Mujeres 25–45 años, < Preparatoria; Grupo 2 = Hombres 25–45 años, < Preparatoria; Grupo 3 = Mujeres 25–45 años, ≥ Preparatoria; Grupo 4 = Hombres 25–45 años, ≥ Preparatoria; Grupo 5 = Mujeres 46–65 años, < Preparatoria; Grupo 6 = Hombres 46–65 años, < Preparatoria; Grupo 7 = Mujeres 46–65 años, ≥ Preparatoria; Grupo 8 = Hombres 46–65 años, ≥ Preparatoria.

Fuente: Elaboración propia con datos de la ENOE. Proceso disponible en:  danifuentesga

## 1.9

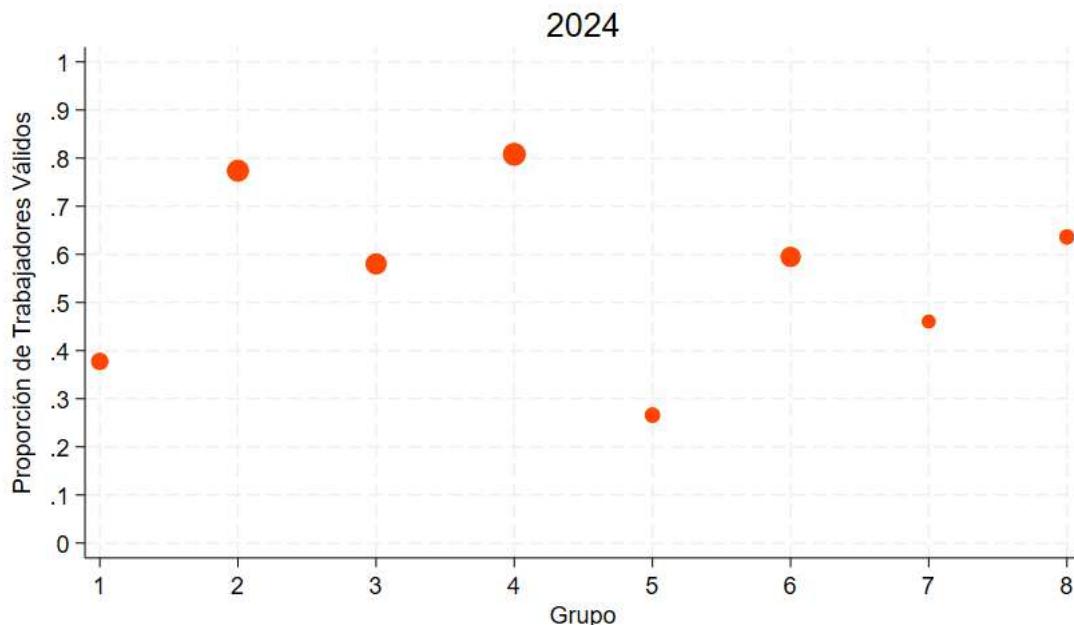
Analiza cómo cambia la proporción de trabajadores para la población y para cada grupo (*asegúrate de utilizar los pesos factor que incluye la ENIGH y ENOE*).

Utiliza el comando `table`. Discute qué grupo ha sido más afectado o beneficiado en términos de participación laboral.

### RESPUESTA:

#### 1.9.1 Resultados ENIGH

Figura 21: Distribución de trabajadores válidos por grupo poblacional en la ENIGH, 2024)



Fuente: Elaboración propia con datos de la ENIGH. Proceso disponible en:  danifuentesga

La figura muestra la **proporción de trabajadores válidos** en cada grupo poblacional para el año 2024. Se observa que los grupos con **mayor nivel educativo** (como el 3 y el 4) alcanzan las tasas más altas de participación laboral, mientras que en los grupos con menor escolaridad y mayor edad (como el 5 y el 7) la proporción es considerablemente menor. Esto

refleja cómo la **educación y la edad influyen directamente en la inserción laboral**, mostrando una ventaja clara para las personas más jóvenes y con mayor escolaridad.

Al revisar la **participación laboral** entre 1992 y 2024 de la siguiente tabla, se observa que los grupos con mayor **beneficio** han sido los de **mujeres con preparatoria o más** (Grupos 3 y 7), cuya proporción de trabajadoras aumentó de manera sostenida. En contraste, los grupos más **afectados** son los de **hombres con baja escolaridad** (Grupos 2 y 6), que muestran una caída o estancamiento en su participación laboral a lo largo del periodo.

La **proporción de trabajadores con ingreso válido** muestra un aumento claro entre 1992 y 2024: pasó de alrededor del **40 %** a más del **56 %**. El crecimiento fue especialmente marcado a partir de los **años dos mil**, cuando se superó por primera vez la barrera del 50 %. Esto refleja una mayor **formalización y participación** en el mercado laboral con ingresos reportados, aunque con algunas fluctuaciones en periodos recientes.

**Tabla 22.** Proporción de trabajadores en la población por Grupo Poblacional en la ENIGH, 1992–2024

Año	Grupo 1	Grupo 2	Grupo 3	Grupo 4	Grupo 5	Grupo 6	Grupo 7	Grupo 8
1992	19.88	67.05	55.93	74.75	9.28	41.32	36.34	65.03
1994	20.41	64.89	55.66	74.07	9.40	40.99	36.38	69.23
1996	22.64	66.20	53.73	74.28	11.49	39.99	41.50	63.73
1998	22.49	68.75	51.17	66.86	12.31	45.83	40.91	53.25
2000	24.98	71.03	48.43	68.31	13.49	47.13	32.35	55.87
2002	27.23	70.10	52.97	66.71	15.41	48.70	39.37	59.16
2004	28.19	75.73	53.01	77.53	15.20	49.72	38.67	57.50
2005	26.26	73.98	54.36	74.95	15.92	51.27	39.96	57.55
2006	28.66	73.09	54.43	77.55	17.18	51.05	41.56	61.02
2008	29.84	72.58	56.30	79.19	20.33	51.38	38.54	63.39
2010	30.45	72.47	54.18	76.29	19.96	50.60	43.02	62.92
2012	32.35	70.65	57.70	76.52	21.77	50.49	39.15	58.29
2014	33.71	75.48	53.80	79.69	21.82	53.26	41.07	63.85
2016	37.28	78.02	58.24	80.39	26.00	56.69	42.32	63.71
2018	37.15	77.49	57.20	79.79	26.58	58.40	45.81	63.93
2020	34.93	75.11	53.55	75.99	23.31	54.94	42.01	59.35
2022	39.06	77.07	58.18	79.84	26.94	58.84	43.59	63.63
2024	38.97	76.94	59.29	80.64	27.94	59.48	46.71	64.14

*Nota:* Los grupos de población se definieron según edad, sexo y nivel educativo: Grupo 1 = Mujeres 25–45 años, < Preparatoria; Grupo 2 = Hombres 25–45 años, < Preparatoria; Grupo 3 = Mujeres 25–45 años, ≥ Preparatoria; Grupo 4 = Hombres 25–45 años, ≥ Preparatoria; Grupo 5 = Mujeres 46–65 años, < Preparatoria; Grupo 6 = Hombres 46–65 años, < Preparatoria; Grupo 7 = Mujeres 46–65 años, ≥ Preparatoria; Grupo 8 = Hombres 46–65 años, ≥ Preparatoria.

Fuente: Elaboración propia con datos de la ENIGH. Proceso disponible en:  danifuentesga

**Tabla 23.** Proporción de trabajadores en la población con ingreso valido en la ENIGH, 1992–2024

Año	Proporción de trabajadores (%)
1992	40.72
1994	40.43
1996	41.79
1998	41.96
2000	43.55
2002	44.66
2004	48.33
2005	47.78
2006	48.92
2008	50.04
2010	49.68
2012	50.10
2014	52.20
2016	55.01
2018	55.29
2020	52.02
2022	55.43
2024	56.54

Fuente: Elaboración propia con datos de la ENIGH. Proceso disponible en:  danifuentesga

### 1.9.2 Resultados ENOE

**Tabla 24.** Proporción de trabajadores en la población por Grupo, ENOE (2005–2025)

Año	Grupo 1	Grupo 2	Grupo 3	Grupo 4	Grupo 5	Grupo 6	Grupo 7	Grupo 8
2005	32.55	77.80	49.52	74.28	26.66	65.55	37.15	62.12
2006	33.52	77.61	50.12	73.43	27.14	66.28	38.41	61.17
2007	33.45	76.92	48.68	71.10	27.90	64.39	38.28	60.22
2008	33.90	75.42	47.04	68.95	27.59	62.11	36.59	57.35
2009	33.63	72.49	45.52	63.89	28.10	60.34	33.85	53.83
2010	34.17	72.67	44.74	64.46	28.37	59.90	34.21	52.37
2011	33.91	71.74	42.40	61.60	28.68	59.37	32.88	49.60
2012	35.15	71.69	43.40	61.08	29.92	59.17	32.39	50.42
2013	35.09	70.18	41.48	59.41	29.26	57.22	31.33	47.61
2014	34.17	70.97	39.90	57.71	28.72	57.75	30.19	46.81
2015	34.39	70.49	38.96	56.07	29.31	57.51	30.46	45.21
2016	34.73	70.00	38.33	56.12	29.11	57.11	29.95	44.56
2017	33.59	68.57	36.82	53.61	28.05	56.28	28.33	43.06
2018	33.57	67.80	36.40	51.93	28.63	55.90	27.84	42.51
2019	35.39	67.89	37.90	53.76	30.13	56.82	29.50	44.68
2020	33.87	66.19	37.86	52.30	27.78	54.17	28.05	42.88
2021	33.91	66.47	38.63	53.04	27.88	54.30	28.56	43.20
2022	35.56	66.75	39.18	53.33	29.35	56.13	29.91	43.33
2023	35.79	65.87	39.29	52.67	31.13	55.83	29.87	43.01
2024	35.36	64.22	38.91	51.52	29.80	54.59	29.86	43.19
2025	32.99	59.05	34.70	45.83	27.59	49.35	25.49	35.97

*Nota:* Los grupos de población se definieron según edad, sexo y nivel educativo: Grupo 1 = Mujeres 25–45 años, < Preparatoria; Grupo 2 = Hombres 25–45 años, < Preparatoria; Grupo 3 = Mujeres 25–45 años, ≥ Preparatoria; Grupo 4 = Hombres 25–45 años, ≥ Preparatoria; Grupo 5 = Mujeres 46–65 años, < Preparatoria; Grupo 6 = Hombres 46–65 años, < Preparatoria; Grupo 7 = Mujeres 46–65 años, ≥ Preparatoria; Grupo 8 = Hombres 46–65 años, ≥ Preparatoria.

Fuente: Elaboración propia con datos de la ENOE. Proceso disponible en:  danifuentesga

**Tabla 25.** Proporción de trabajadores en la población con ingreso valido en la ENOE, 2005-2025

Año	Proporción de trabajadores (%)
2005	52.41
2006	52.79
2007	51.94
2008	50.89
2009	49.07
2010	49.17
2011	48.06
2012	48.44
2013	47.17
2014	46.51
2015	46.06
2016	45.72
2017	44.15
2018	43.59
2019	44.90
2020	43.20
2021	43.41
2022	44.24
2023	44.16
2024	43.29
2025	38.84

## 1.10

¿Qué tan comparable es el ingreso laboral en ENIGH 2018 a los anteriores?

¿Qué grupos fueron más afectados?

Leer de internet la discusión de la comparabilidad de ingreso con la nueva **ENIGH**.

### **RESPUESTA:**

La **ENIGH 2018** introdujo cambios metodológicos importantes en la captación del ingreso laboral: nuevas preguntas sobre **prestaciones**, ajustes en los rubros de **servicios médicos, licencias e incapacidades** y modificaciones en el registro del **ahorro para el retiro**. Estas innovaciones mejoraron la **desagregación estadística**, pero generaron dudas sobre la **comparabilidad histórica** con ediciones previas. Instituciones como el **CONEVAL** señalaron la necesidad de aplicar el **Modelo Estadístico de Continuidad (MEC)** para suavizar series y evitar saltos artificiales, mientras que investigadores del **PUED** y del **CEEY** advirtieron en foros y redes sociales que algunas alzas en los ingresos reportados podrían reflejar más una **ilusión estadística** que un verdadero crecimiento en el poder adquisitivo.

En cuanto a los **grupos más afectados**, la evidencia apunta a que los cambios golpearon sobre todo a los hogares con menor escolaridad y a los trabajadores en la **informalidad**, pues son quienes reportan con mayor dificultad ingresos no salariales o irregulares. Al mismo tiempo, los grupos con **educación media y superior**, vinculados al empleo formal, parecen haber sido los más “beneficiados” de la nueva captación: sus salarios aparecen más estables e incluso con incrementos entre 2016 y 2018. En conjunto, el debate académico en torno a la ENIGH 2018 muestra que la **comparabilidad del ingreso** debe leerse con cautela, ya que parte de las brechas observadas responden tanto a **cambios reales en el mercado laboral** como a **ajustes en la forma de medirlo**.

## Problema 2: IMSS

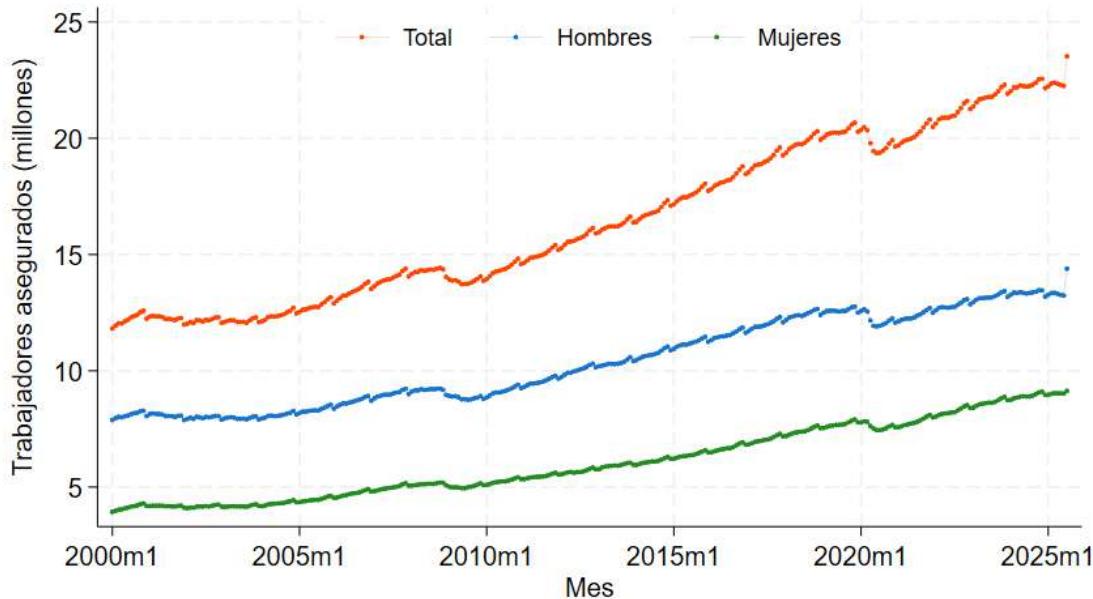
El **objetivo** es tener características de individuos y salarios para diferentes puntos en el tiempo

### 2.1

Usando los microdatos del **IMSS** para todos los meses, grafica el número de empleados totales en el **IMSS** (*solo celdas con masa salarial positiva*), por hombre y por mujer.

#### RESPUESTA:

Figura 22: Trabajadores por sexo registrados en el IMSS, 2000-2025



Fuente: Elaboración propia con datos del IMSS. Proceso disponible en:  danifuentesga

### 2.2

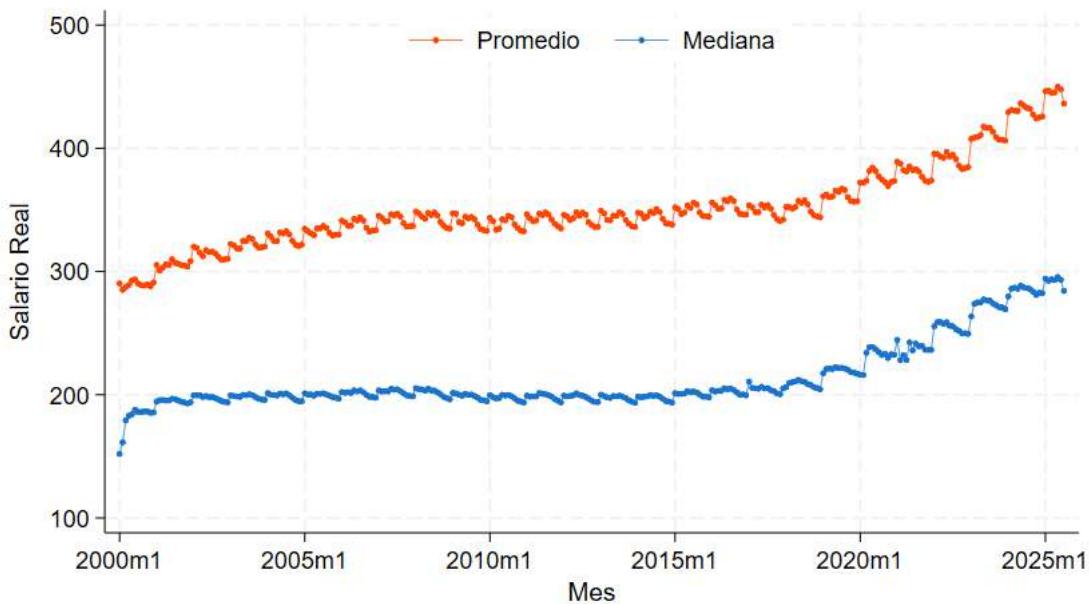
Pon los **ingresos reales** en el mismo periodo que el problema 1. Lleva los ingresos a **ingresos mensuales**.

Grafica el promedio de los ingresos y la mediana de los ingresos para todos los trabajadores en todo el periodo. Razoná cómo realizar esto pues tienes celdas a nivel de unidades (*diferentes combinaciones de celdas*).

**Hint:** calcula el **salario promedio** (*masa sobre número de trabajadores*), y luego usa el número de trabajadores como ponderador.

### RESPUESTA:

Figura 23: Promedio y Mediana de ingresos diarios registrados en el IMSS, 2000-2025



Fuente: Elaboración propia con datos del IMSS. Proceso disponible en:  danifuentesga

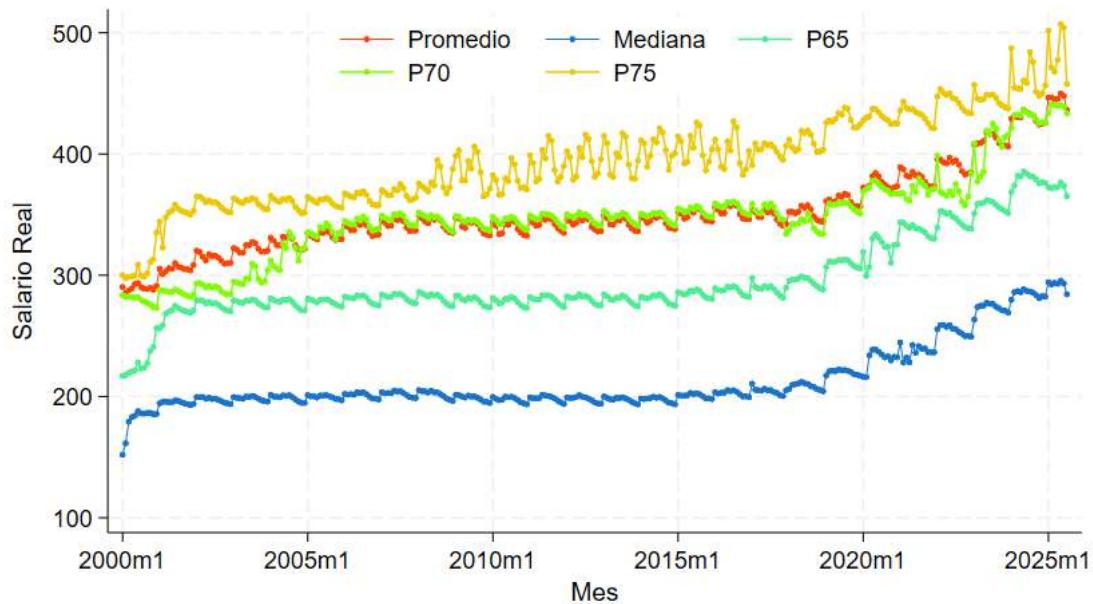
## 2.3

Grafica el **promedio** y la **mediana del salario al mes** para todos los trabajadores.

¿En qué percentil se encuentra el promedio?

**RESPUESTA:** El promedio se acerca más a los percentiles alto de la distribución , específicamente p65 y p75

Figura 24: Promedio y Mediana por percentil del ingreso diario con datos del IMSS, 2000-2025



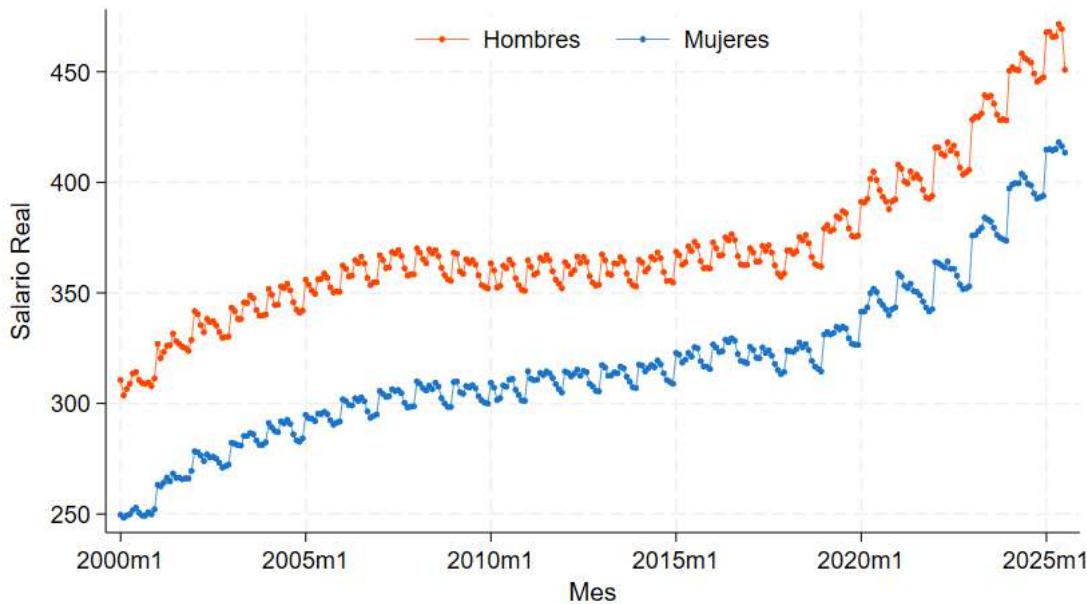
Fuente: Elaboración propia con datos del IMSS. Proceso disponible en:  danifuentesga

## 2.4

Grafica la **brecha salarial de género** para todo el periodo.

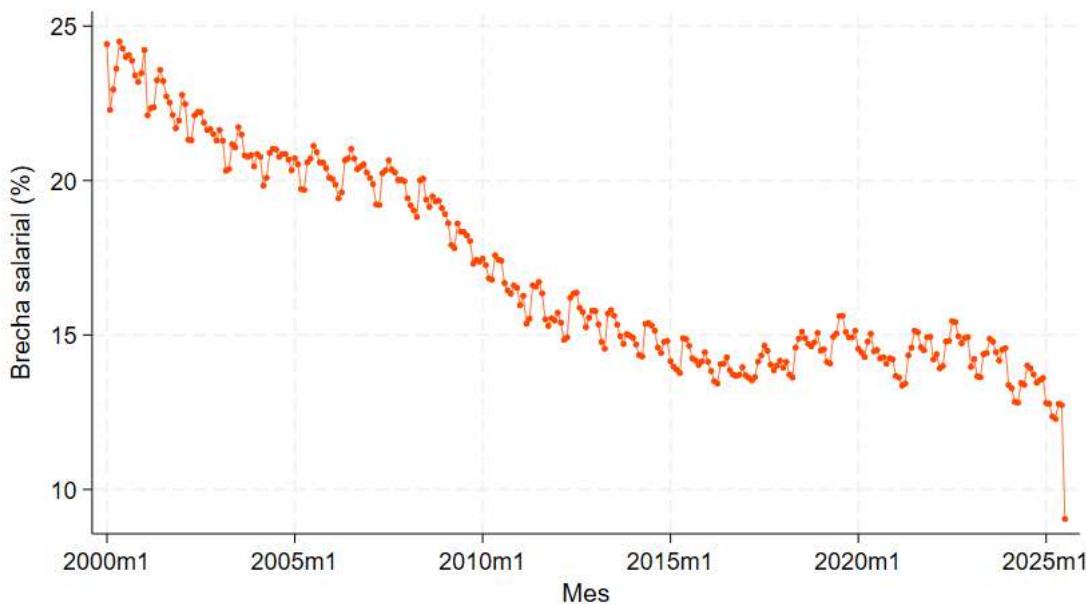
**RESPUESTA:**

Figura 25: Brecha Salarial (Diario) por Género con datos del IMSS, 2000-2025



Fuente: Elaboración propia con datos del IMSS. Proceso disponible en:  danifuentesga

Figura 26: Tamaño de la Brecha Salarial por Género con datos del IMSS, 2000-2025



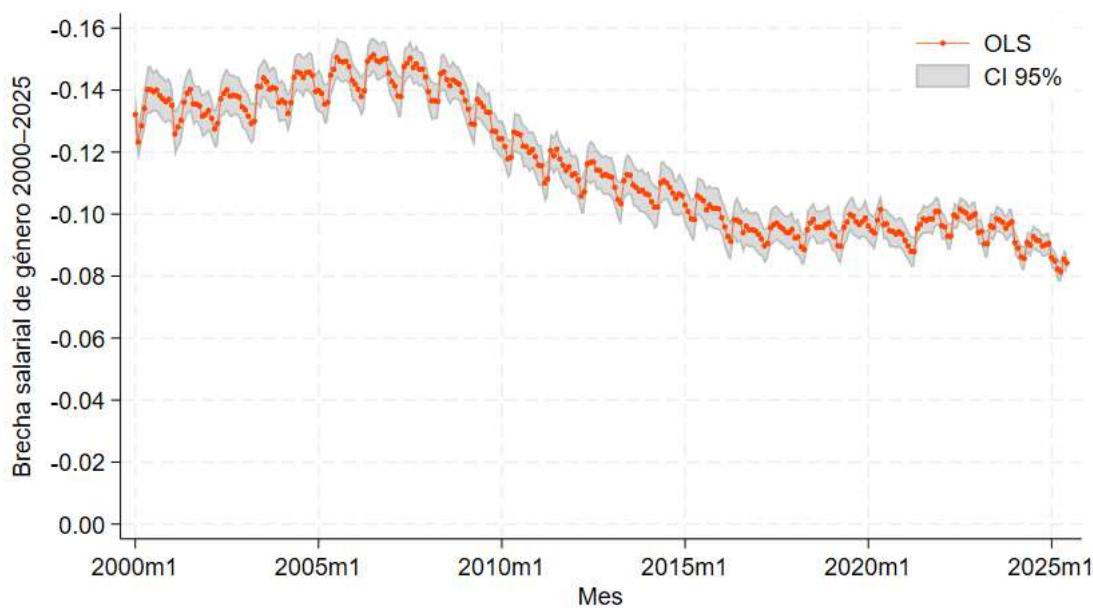
Fuente: Elaboración propia con datos del IMSS. Proceso disponible en:  danifuentesga

## 2.5

Realiza un **bootstrap** para obtener **intervalos de confianza** con el *método del percentil 95 %*, e inclúyelos en tu figura de **brecha de género**.

### RESPUESTA:

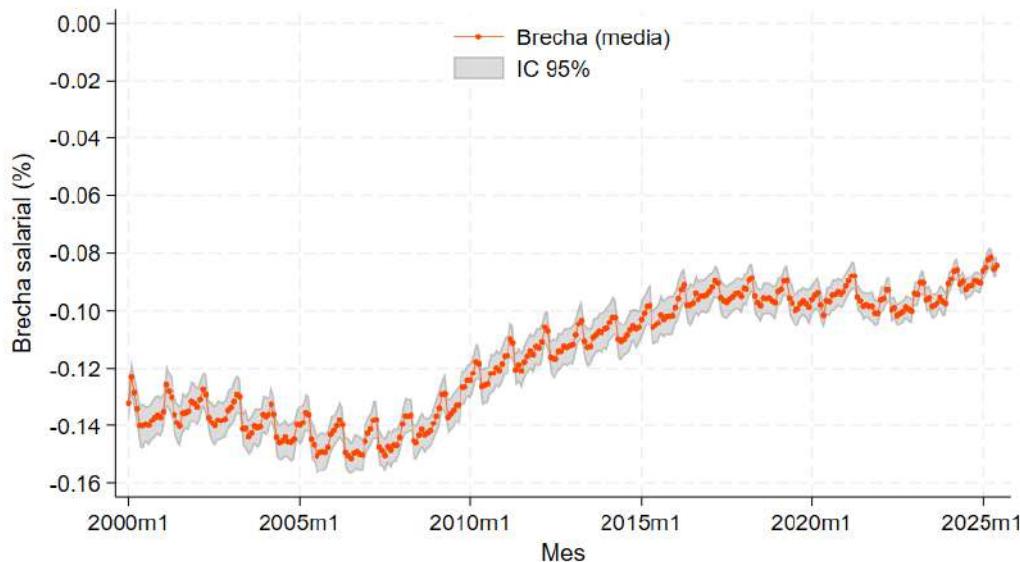
Figura 27: Brecha salarial de género estimada por OLS con intervalos de confianza bootstrap (2000–2025)



Fuente: Elaboración propia con datos del IMSS. Proceso disponible en:  danifuentesga

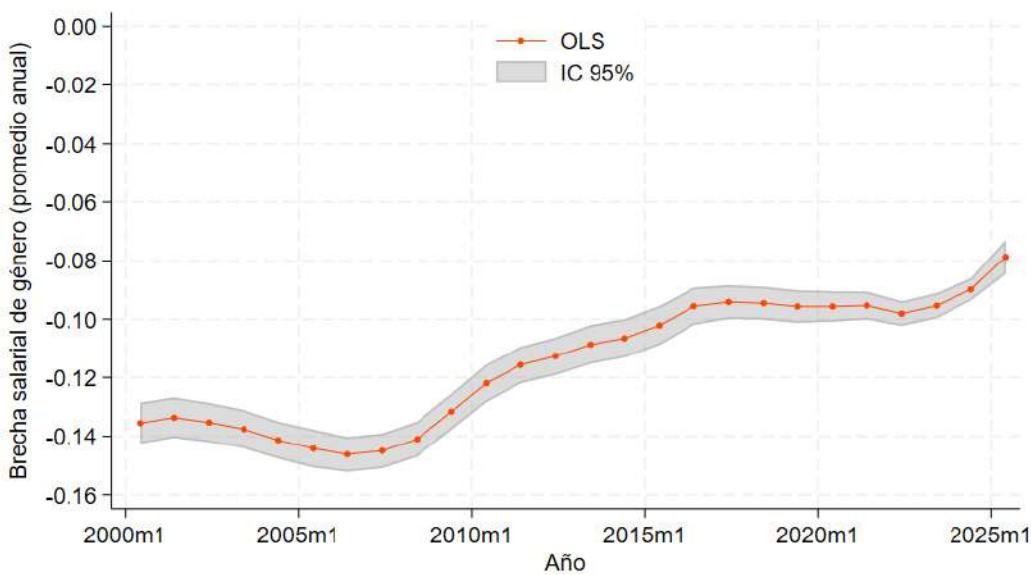
La **brecha salarial de género** se estimó mediante **regresiones OLS** para cada mes entre 2000 y 2025, incorporando **intervalos de confianza** obtenidos con un procedimiento **bootstrap al 95 %**. La gráfica muestra que la brecha pasó de valores cercanos a **-0.14** en los primeros años a alrededor de **-0.09** en la etapa más reciente, con una tendencia a la reducción aunque sin desaparecer, lo que refleja un avance parcial en la **igualdad salarial**.

Figura 28: 2 Brecha salarial de género estimada por OLS con intervalos de confianza bootstrap (2000–2025)



Fuente: Elaboración propia con datos del IMSS. Proceso disponible en:  danifuentesga

Figura 29: Brecha salarial de género en México (2000–2025): promedio anual



Fuente: Elaboración propia con datos del IMSS. Proceso disponible en:  danifuentesga

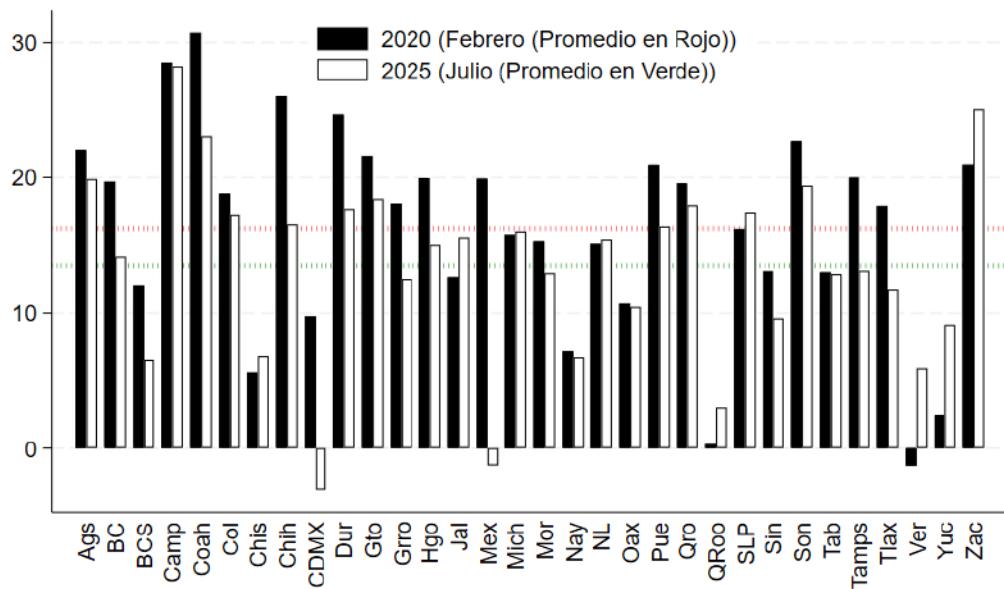
## 2.6

¿Qué entidad federativa tiene la menor y mayor brecha de género en 2025?

¿Qué entidades son las más afectadas por la COVID (*comparar Febrero 2020 con Julio 2025*)?

### RESPUESTA:

Figura 30: Brecha Salarial de Género por Entidad Federativa: Comparación 2020 (Febrero) vs 2025 (Julio)



Fuente: Elaboración propia con datos del IMSS. Proceso disponible en:  danifuentesga

En el año **2025**, la entidad federativa con la **mayor brecha de género** es **Zacatecas** (**25.1 %**), mientras que la de **menor brecha** es la **Ciudad de México** (**-3.1 %**, lo cual indica que las mujeres, en promedio, superan a los hombres en ingresos formales). Al comparar **febrero de 2020** con **julio de 2025**, se observa que las entidades más afectadas por la **COVID-19** en términos de ampliación de la brecha salarial son **Zacatecas** y **Campeche**, pues muestran incrementos elevados en la brecha, mientras que estados como **Ciudad de México** y **Estado de México** destacan por una reducción considerable de la brecha en el

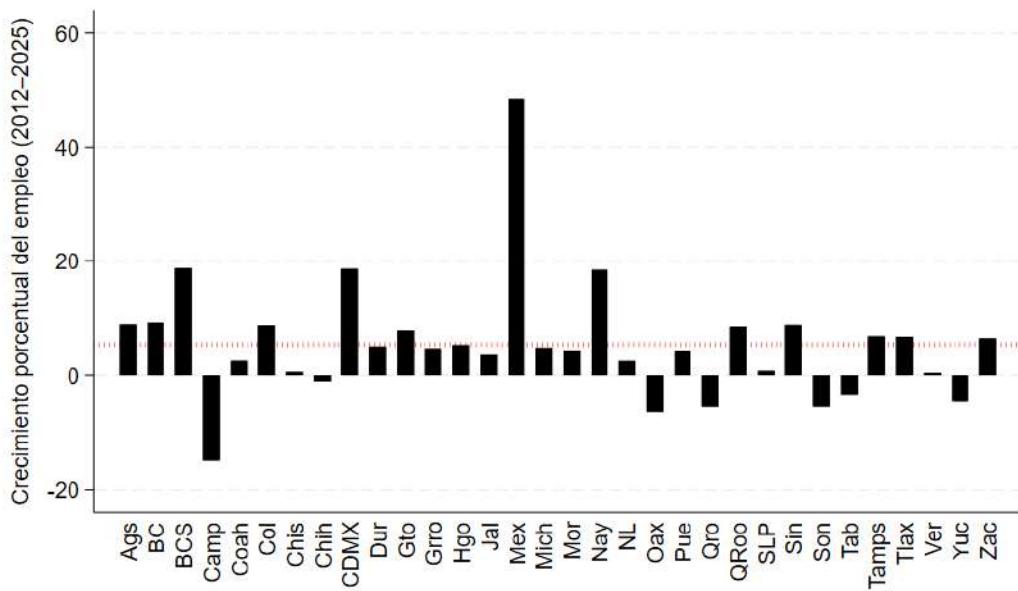
mismo periodo.

## 2.7

¿Qué entidad federativa ha incrementado más su empleo en términos porcentuales entre 2012 y 2025 (*usa el mismo mes o trimestre*)?

### RESPUESTA:

Figura 31: Crecimiento porcentual del empleo formal por entidad federativa (2012–2025)



Fuente: Elaboración propia con datos del IMSS. Proceso disponible en:  danifuentesga

La gráfica muestra el **crecimiento porcentual del empleo formal por entidad federativa entre 2012 y 2025**. En promedio, el empleo aumentó ligeramente en el país, como lo indica la línea roja punteada. Sin embargo, se observan diferencias importantes: **Jalisco** destaca con el mayor crecimiento (superior al 45 %), mientras que **Baja California Sur** y **Campeche** presentan retrocesos significativos, con caídas de hasta -15 %. También se aprecia que entidades como la **Ciudad de México** y **Quintana Roo** mantienen niveles de empleo relativamente altos en ambos años, lo cual refleja un mercado laboral más dinámico.

En contraste, estados como **Tlaxcala** y **Zacatecas** muestran incrementos modestos que se sitúan por debajo del promedio nacional.

## 2.8

Realiza un **mapa en Stata** y en **R** sobre los cambios en empleo y en salario a nivel entidad federativa.

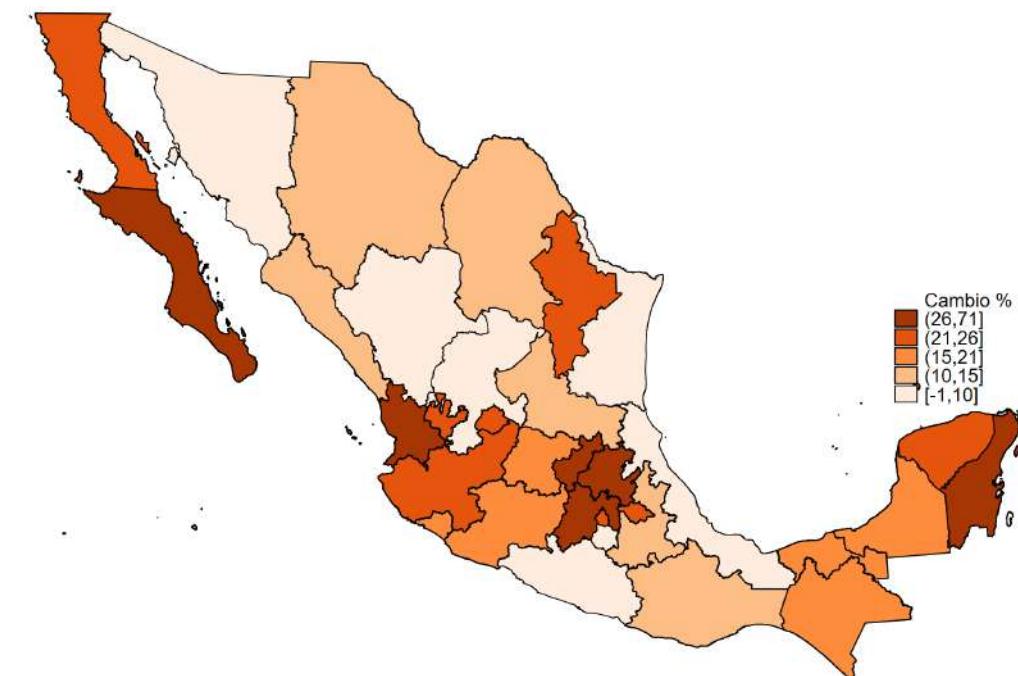
Usa dos años base: **2010** y **2017**, y el año final **2025**. Queremos saber la distribución de cambios en empleo y salario en ese periodo.

### RESPUESTA:

#### 2.8.1 Mapas en STATA

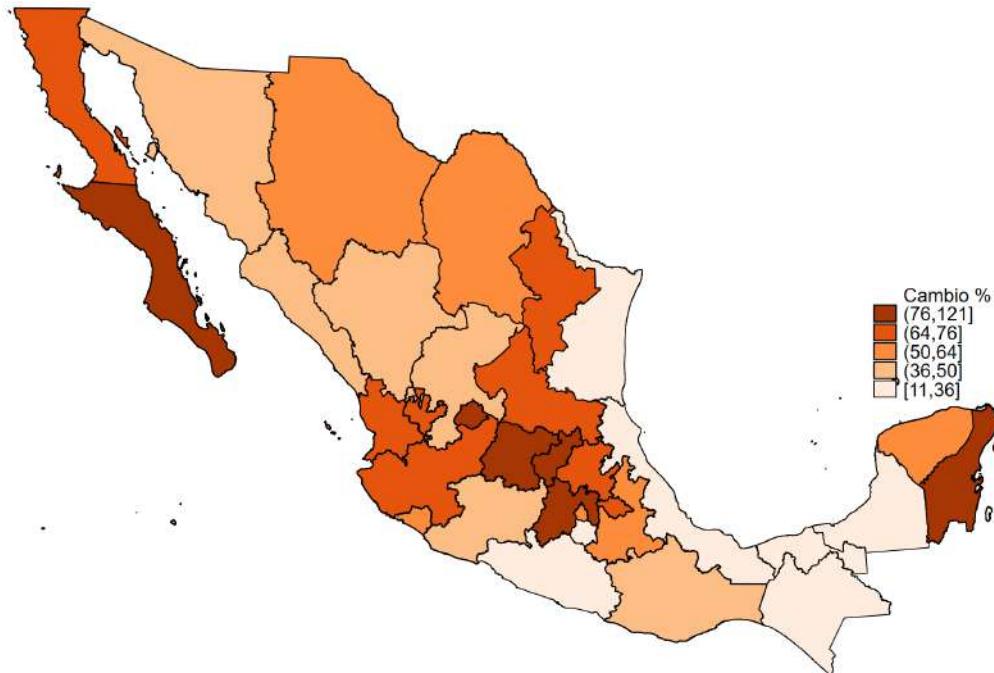
##### STATA

Figura 32: Crecimiento porcentual del empleo formal por entidad federativa 2017-2025



Fuente: Elaboración propia con datos del IMSS. Proceso disponible en:  danifuentesga

Figura 33: Crecimiento porcentual del empleo formal por entidad federativa 2010-2025



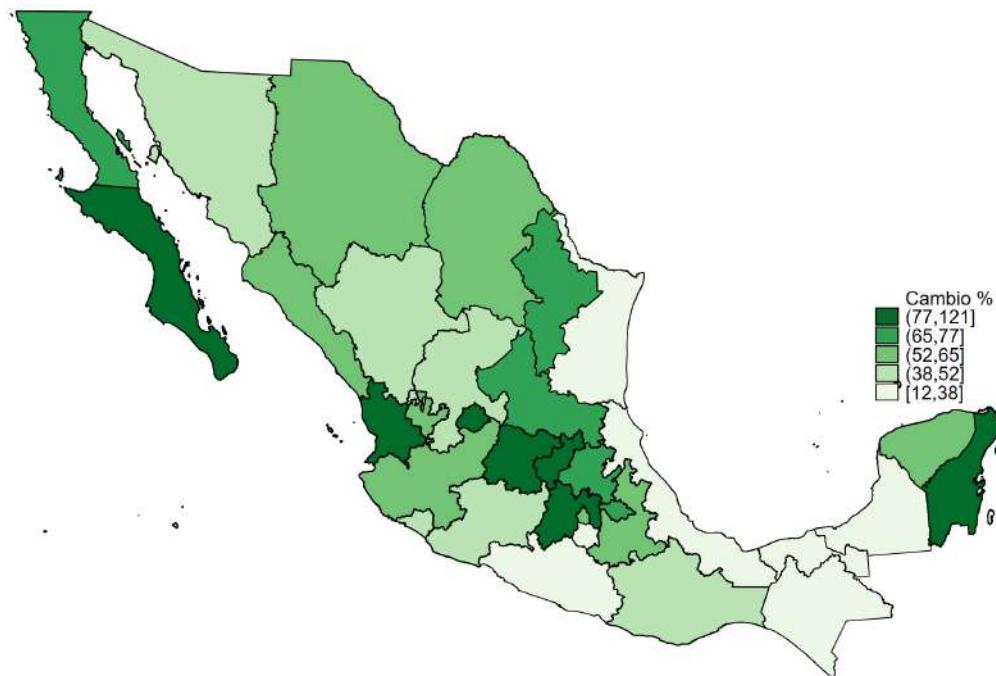
Fuente: Elaboración propia con datos del IMSS. Proceso disponible en:  danifuentesga

Figura 34: Crecimiento porcentual del salario por entidad federativa 2017-2025



Fuente: Elaboración propia con datos del IMSS. Proceso disponible en:  danifuentesga

Figura 35: Crecimiento porcentual del salario por entidad federativa 2010-2025



Fuente: Elaboración propia con datos del IMSS. Proceso disponible en:  danifuentesga

Los mapas muestran el **crecimiento porcentual del empleo formal y de los salarios** por entidad federativa en distintos períodos (2010–2025 y 2017–2025). Se observa una **heterogeneidad regional** marcada: estados del centro y sureste concentran los mayores aumentos, mientras que algunas entidades del norte y noroeste presentan crecimientos más moderados. En general, el empleo formal y los salarios han tenido una **expansión significativa**, aunque con ritmos distintos entre regiones, lo que refleja tanto **dinámicas locales de desarrollo** como diferencias estructurales en los mercados laborales estatales.

### 2.8.1 Mapas en RStudio

RStudio

Figura 36: Crecimiento porcentual del empleo formal por entidad federativa 2017-2025

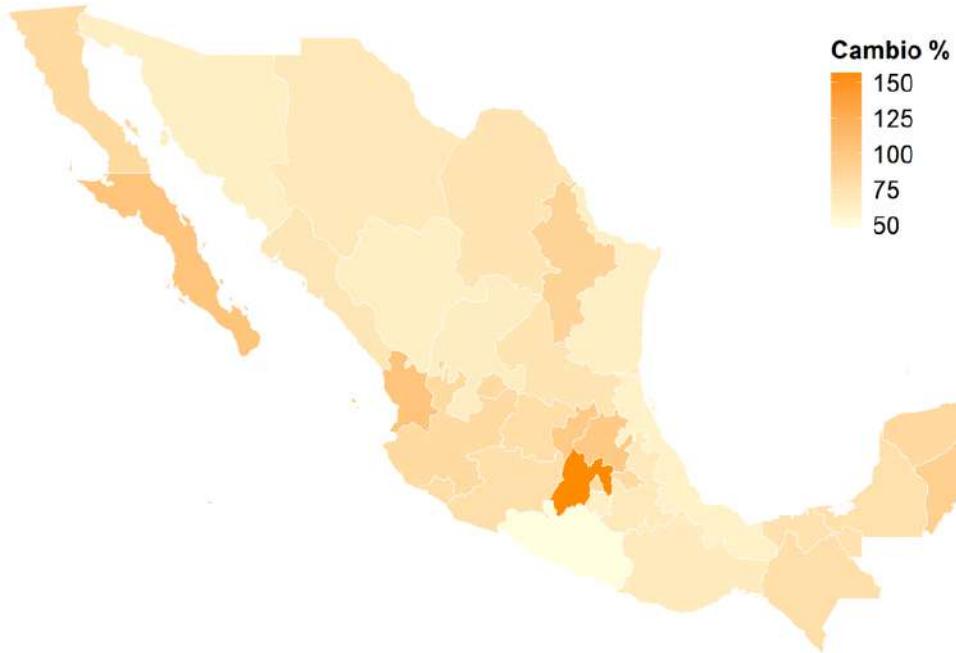


Figura 37: Crecimiento porcentual del empleo formal por entidad federativa 2010-2025

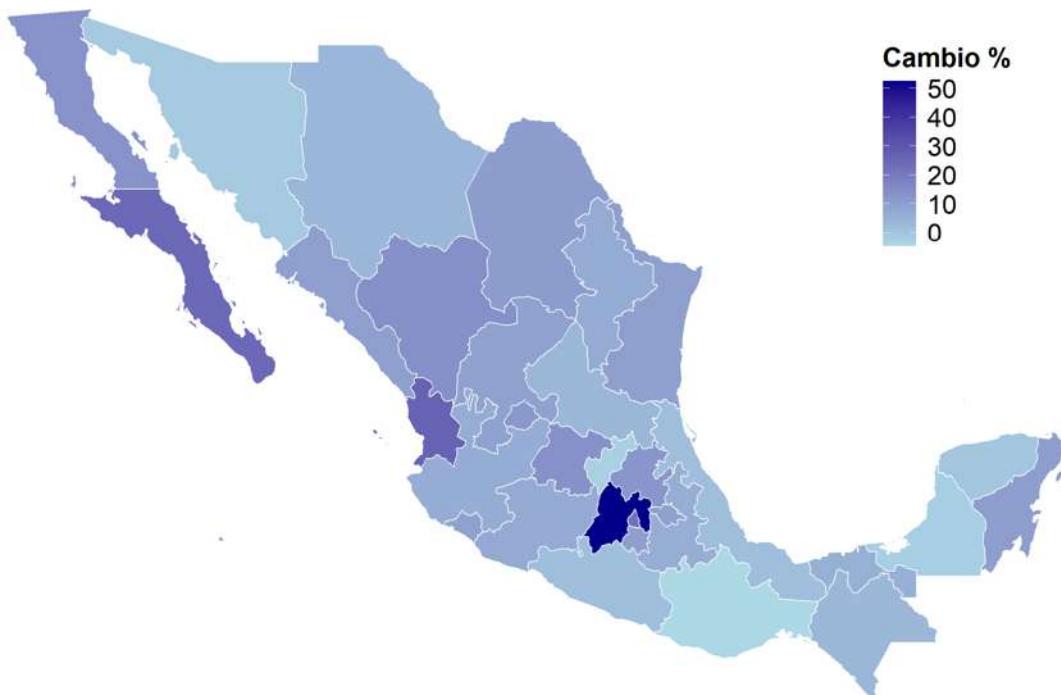


Fuente: Elaboración propia con datos del IMSS. Proceso disponible en:  danifuentesga

Figura 38: Crecimiento porcentual del salario por entidad federativa 2017-2025



Figura 39: Crecimiento porcentual del salario por entidad federativa 2010-2025



Fuente: Elaboración propia con datos del IMSS. Proceso disponible en:  danifuentesga

## 2.9

Realiza un **mapa en Stata** y en **R** sobre los cambios en empleo y en salario a nivel entidad federativa.

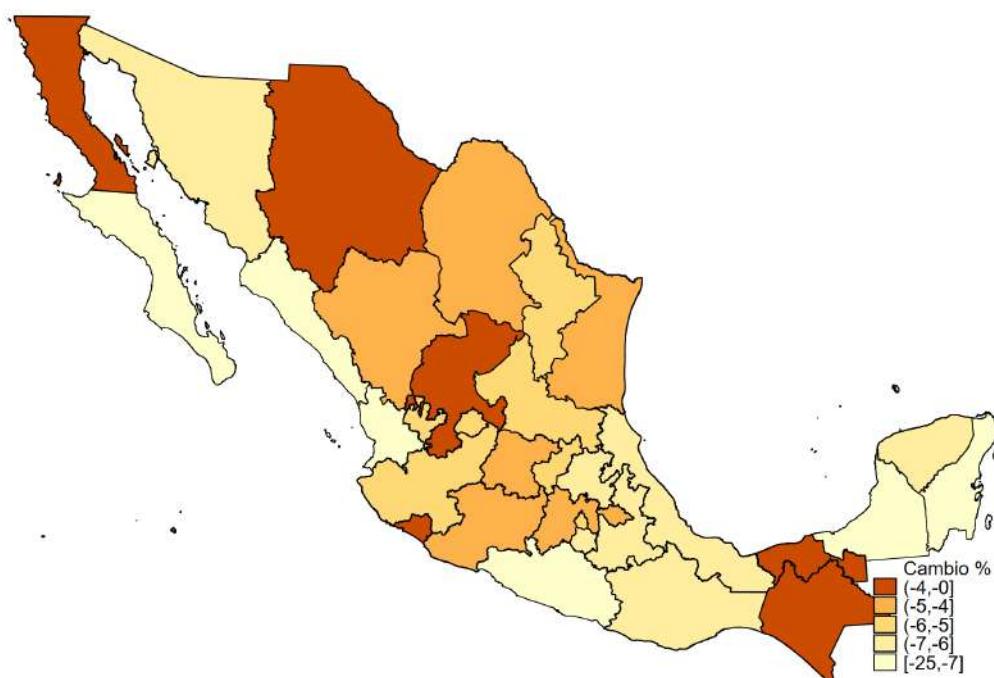
Usa **Febrero 2020** como año base, y como año final **Julio 2020 y 2025**. Queremos saber la **distribución de cambios** en empleo y salario en ese periodo.

### **RESPUESTA:**

#### 2.9.1 Mapas en STATA

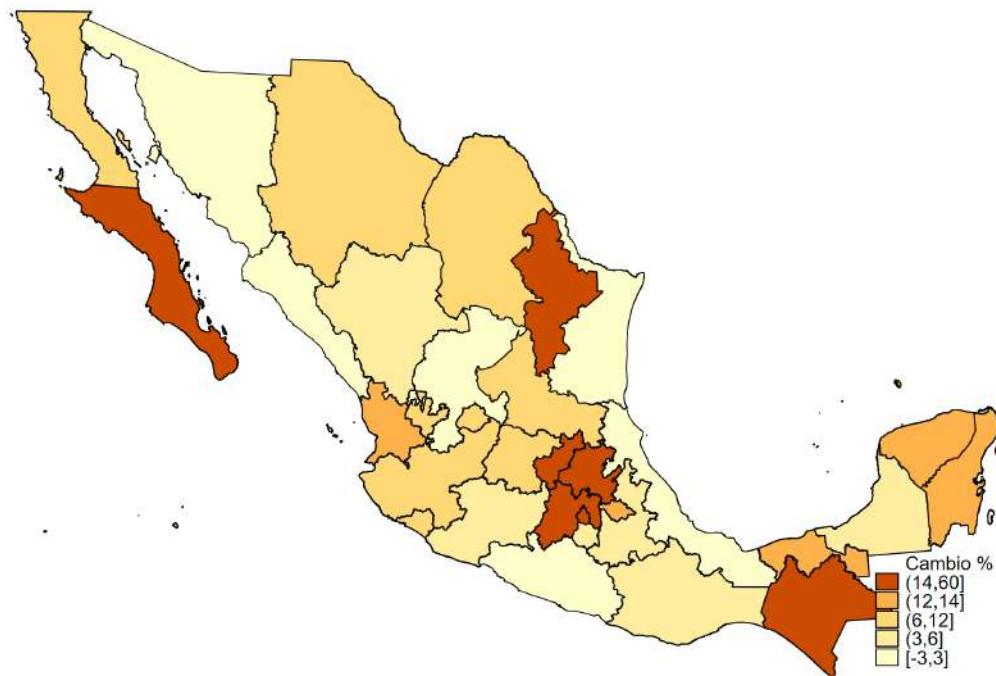
##### STATA

Figura 40: Crecimiento porcentual del empleo por entidad federativa Febrero 2020 - Julio 2020



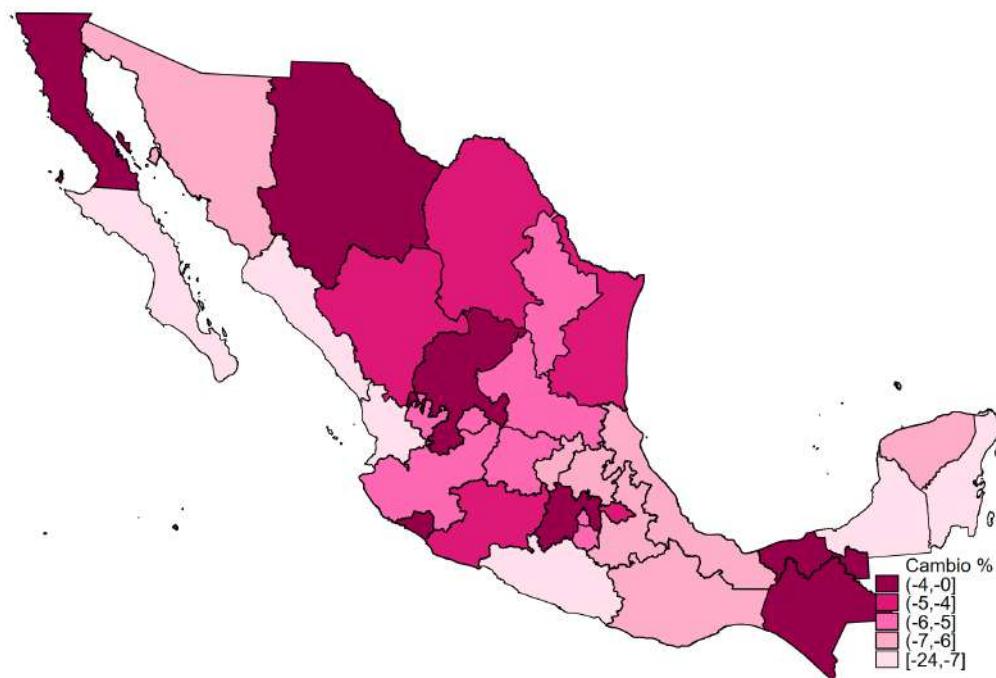
Fuente: Elaboración propia con datos del IMSS. Proceso disponible en:  danifuentesga

Figura 41: Crecimiento porcentual del empleo por entidad federativa Febrero 2020 - Julio 2025



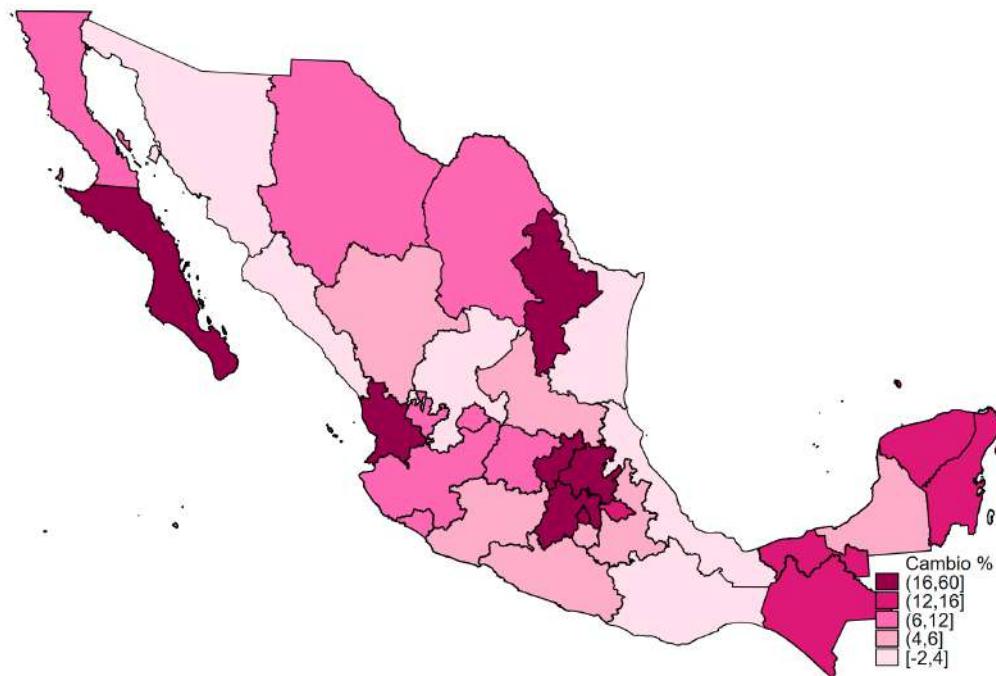
Fuente: Elaboración propia con datos del IMSS. Proceso disponible en:  danifuentesga

Figura 42: Crecimiento porcentual del salario por entidad federativa Febrero 2020 - Julio 2020



Fuente: Elaboración propia con datos del IMSS. Proceso disponible en:  danifuentesga

Figura 43: Crecimiento porcentual del salario por entidad federativa Febrero 2020 - Julio 2025



Fuente: Elaboración propia con datos del IMSS. Proceso disponible en:  danifuentesga

Los mapas reflejan el **impacto regional de la pandemia** en empleo y salarios entre 2020 y 2025. En el corto plazo (2020) casi todas las entidades registraron **caídas en el empleo formal**, con mayor severidad en el norte y sureste. Sin embargo, hacia 2025 varios estados muestran ya **recuperación positiva**, aunque desigual. En paralelo, los salarios tuvieron un **ajuste inicial negativo**, pero después se observa un **crecimiento moderado y más extendido territorialmente**, destacando zonas del Bajío y norte del país. Esto revela que la **recuperación laboral** no fue homogénea: el empleo se recompuso con mayor dificultad que los salarios.

### 2.9.1 Resultados en RStudio

RStudio

Figura 44: Crecimiento porcentual del empleo por entidad federativa Febrero 2020 - Julio 2020



Figura 45: Crecimiento porcentual del empleo por entidad federativa Febrero 2020 - Julio 2025



Fuente: Elaboración propia con datos del IMSS. Proceso disponible en:  danifuentesga

Figura 46: Crecimiento porcentual del salario por entidad federativa Febrero 2020 - Julio 2020



Figura 47: Crecimiento porcentual del salario por entidad federativa Febrero 2020 - Julio 2025



Fuente: Elaboración propia con datos del IMSS. Proceso disponible en:  danifuentesga

## 2.10

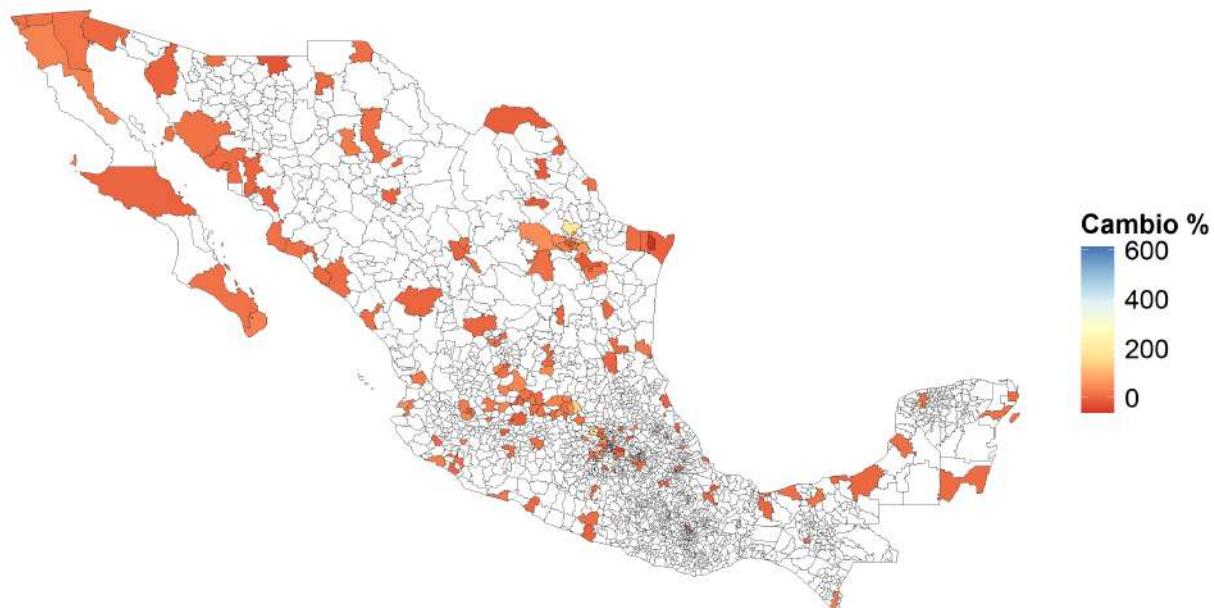
Realiza un **mapa en R** a nivel **municipal** sobre los cambios en empleo y en salario.

Restringe la muestra a aquellos **municipios con al menos 10,000 empleados en el año base**. Para poder hacer este ejercicio, tendrás que hacer un `merge` con la base de municipios de código INEGI (*ver el diccionario que presenta el IMSS*).

Usa como año base **2018**, y como años finales **2019** y **2025**. Queremos saber la **distribución de cambios** en empleo y salario en ese periodo.

### RESPUESTA:

Figura 48: Crecimiento porcentual del empleo por municipio 2018-2019



*Nota:* Solo se consideran municipios que tengan al menos 10,000 empleados en el respectivo año base, lo cual explica mayor presencia de municipios en "blanco".

Fuente: Elaboración propia con datos del IMSS. Proceso disponible en:  danifuentesga

Figura 49: Crecimiento porcentual del empleo por municipio 2018-2025

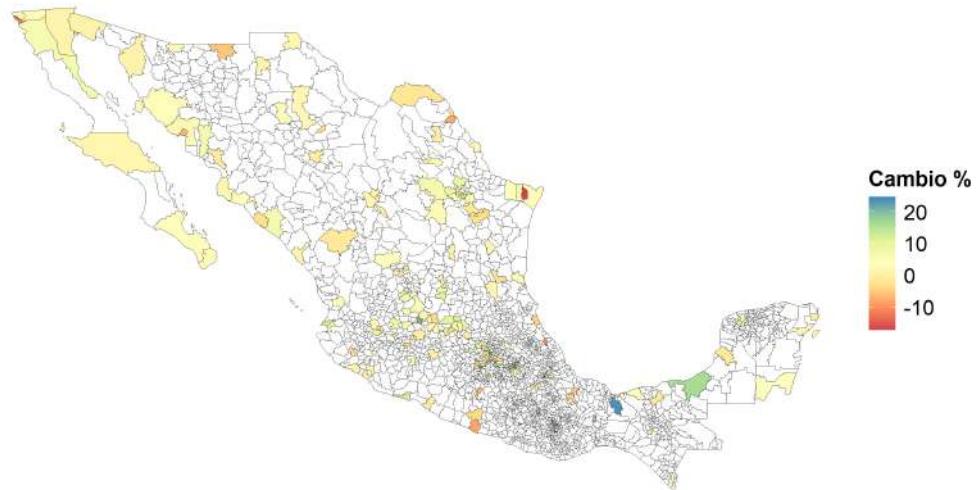
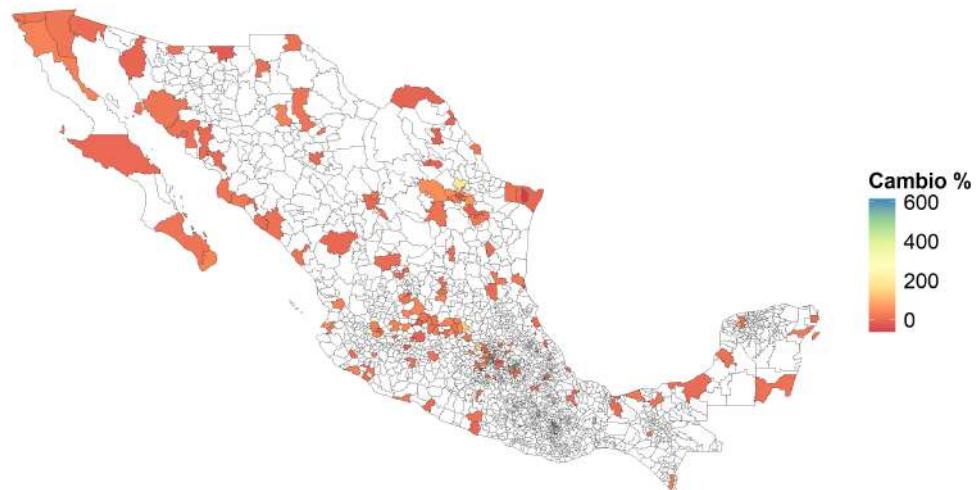


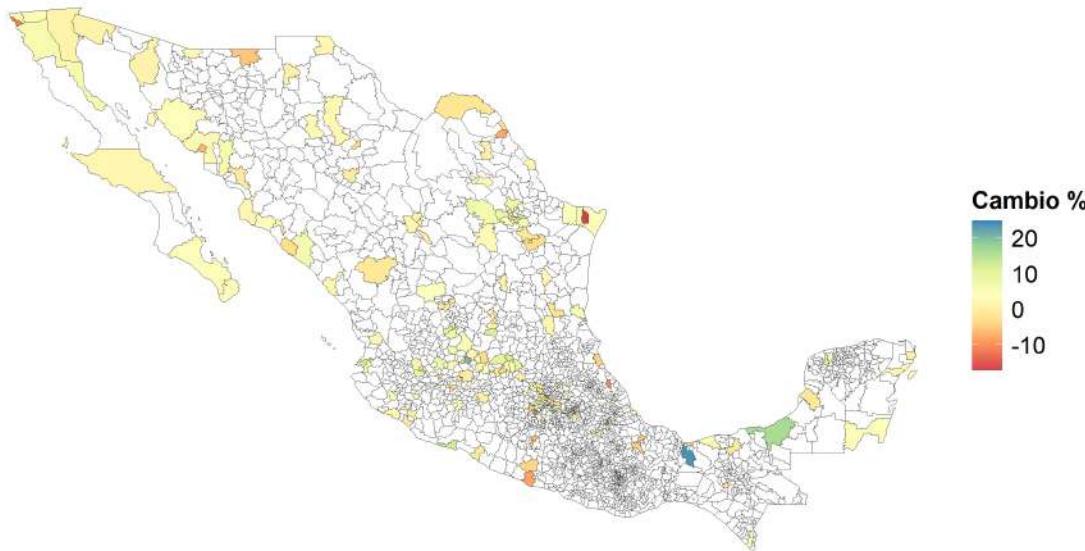
Figura 50: Crecimiento porcentual del salario por municipio 2018-2019



*Nota:* Solo se consideran municipios que tengan al menos 10,000 empleados en el respectivo año base, lo cual explica mayor presencia de municipios en "blanco".

Fuente: Elaboración propia con datos del IMSS. Proceso disponible en:  danifuentesga

Figura 51: Crecimiento porcentual del salario por municipio 2018-2025



*Nota:* Solo se consideran municipios que tengan al menos 10,000 empleados en el respectivo año base, lo cual explica mayor presencia de municipios en "blanco".

Fuente: Elaboración propia con datos del IMSS. Proceso disponible en:  danifuentesga

## 2.11

Con base en tu **análisis anterior**, ¿qué se podría decir sobre el **cambio del salario mínimo en la zona de la frontera norte**? Describe los cambios al respecto.

### **RESPUESTA:**

En la **frontera norte** se observa un **crecimiento porcentual claramente superior** al promedio nacional. Los mapas muestran que entidades como **Baja California, Sonora, Chihuahua, Coahuila, Nuevo León y Tamaulipas** presentan un **aumento sostenido del salario** entre 2020 y 2025, con variaciones que superan el **16%** en varios casos. En contraste, otras regiones del país registran cambios más moderados o incluso **disminuciones**.

Esto confirma que la **zona fronteriza** concentró los **mayores beneficios relativos** en términos de crecimiento salarial en el periodo analizado.

## 2.12

**Control Sintético: Efecto del programa de salario mínimo y reducción de IVA en la ZLFN**

**Objetivo:** Estimar el efecto del programa de **salario mínimo y reducción de IVA** en la **Zona Libre de la Frontera Norte (ZLFN)** a nivel entidad federativa. El programa comienza en **2019**. Se utilizarán los meses de **enero 2015 a julio 2025**.

**Metodología:**

- Realiza un **control sintético a nivel entidad**.
- La unidad tratada será la **ZLFN**, agregando todos los municipios que la conforman.
- Las unidades de control serán las **31 entidades federativas restantes**.
- Realiza un **collapse sum** de empleo y de masa salarial para cada entidad, luego calcula el salario promedio.
- Excluye de los controles:
  - **Baja California** (BC), ya que todos sus municipios pertenecen a la ZLFN.
  - **Tamaulipas**, sólo excluyendo los municipios que forman parte de la ZLFN.

**Se solicita:**

1. Estimar **control sintético** para **10 modelos** (*los mejores en términos de RMSPE*). Discutir cuál se eligió.
2. Estimar el **valor-p** de cada modelo, para las variables de **empleo** y **salario promedio**.
3. Discutir **similitudes o diferencias** con el artículo de **Campos-Vázquez, Delgado y Rodas (2020)** sobre los efectos del salario mínimo en la ZLFN.  
<https://www.sciendo.com/article/10.2478/izajolp-2020-0012>

**RESPUESTA:**

Figura 52: Efectos en la ZLFN en salario promedio (Enero 2015-Julio 2025)

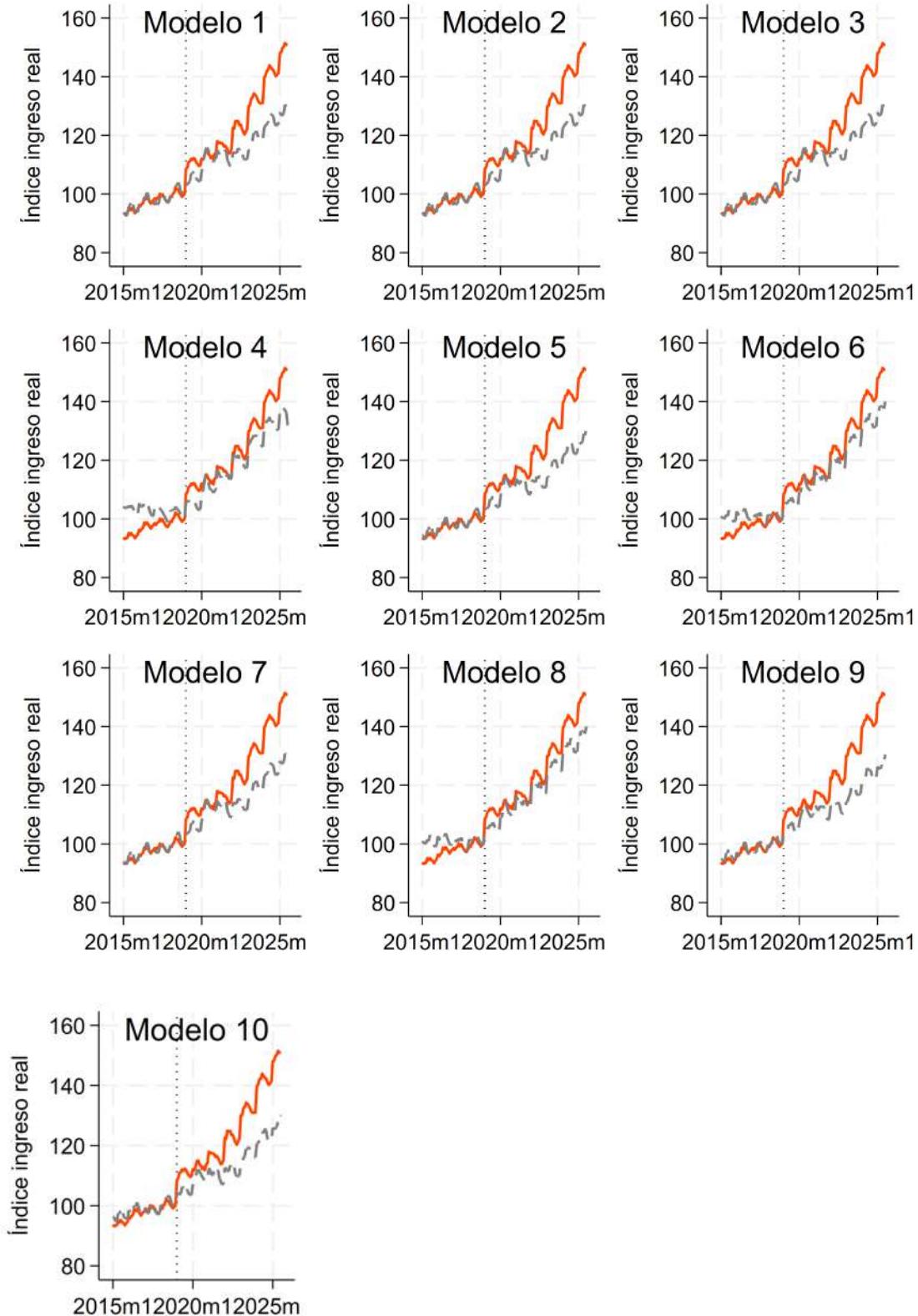


Figura 53: Placebos de modelos de salario promedio (Enero 2015-Julio 2025)

— Tratamiento  
— Donantes (placebos)

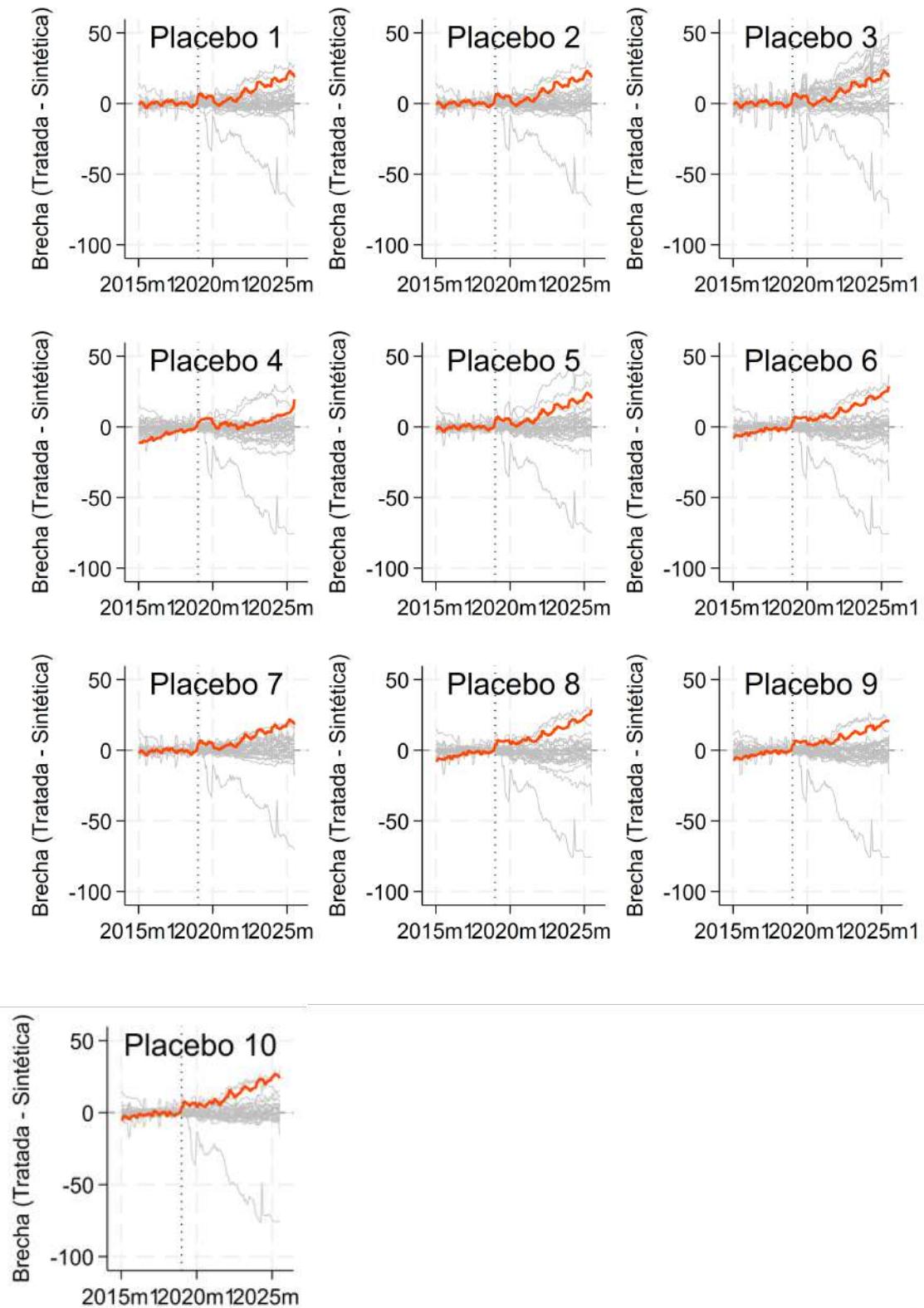
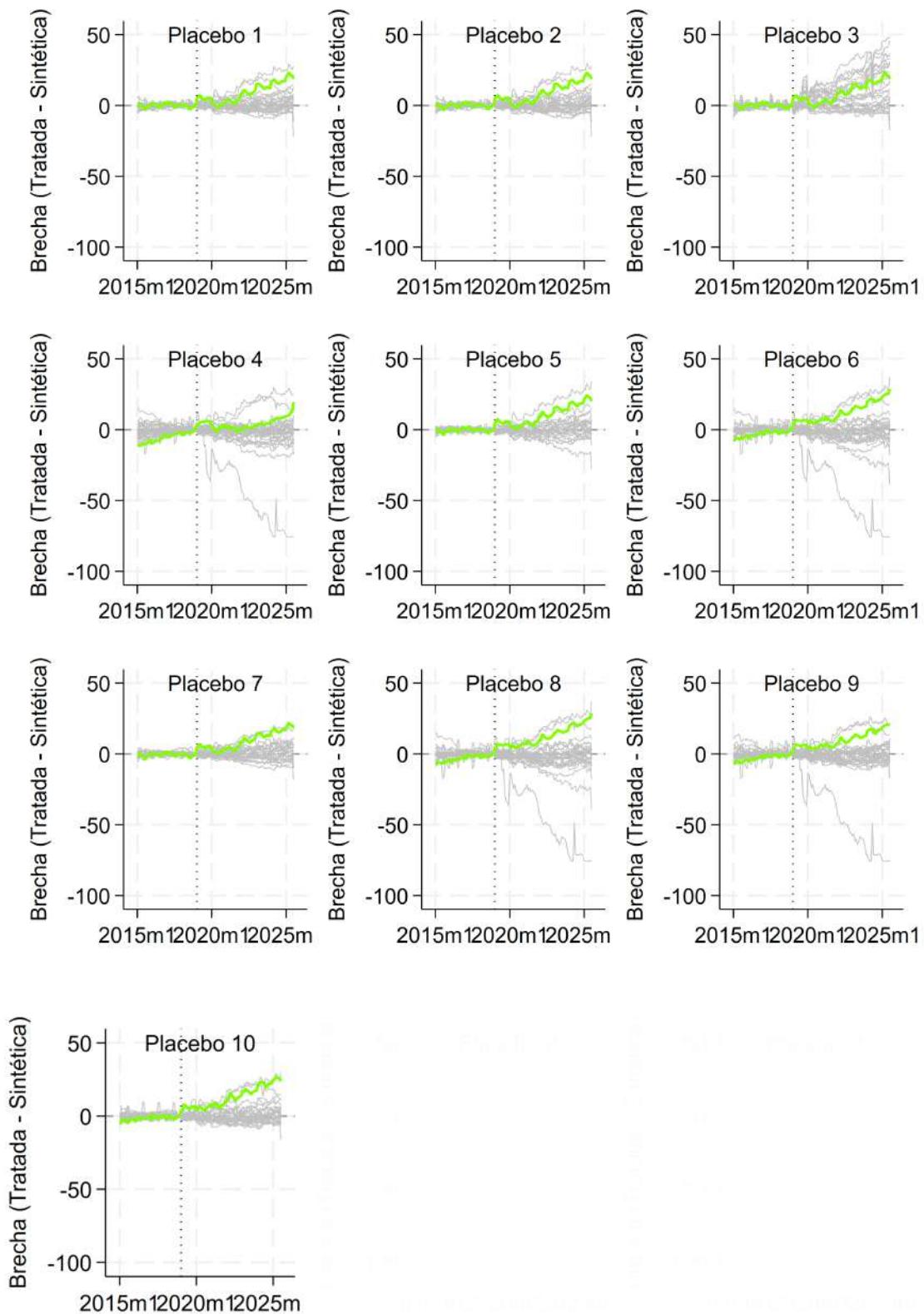


Figura 54: Placebos RMSPE de modelos de salario promedio (Enero 2015-Julio 2025)



Fuente: Elaboración propia con datos del IMSS. Proceso disponible en:  danifuentesga

Figura 55: Efectos en la ZLFN en empleo (Enero 2015-Julio 2025)

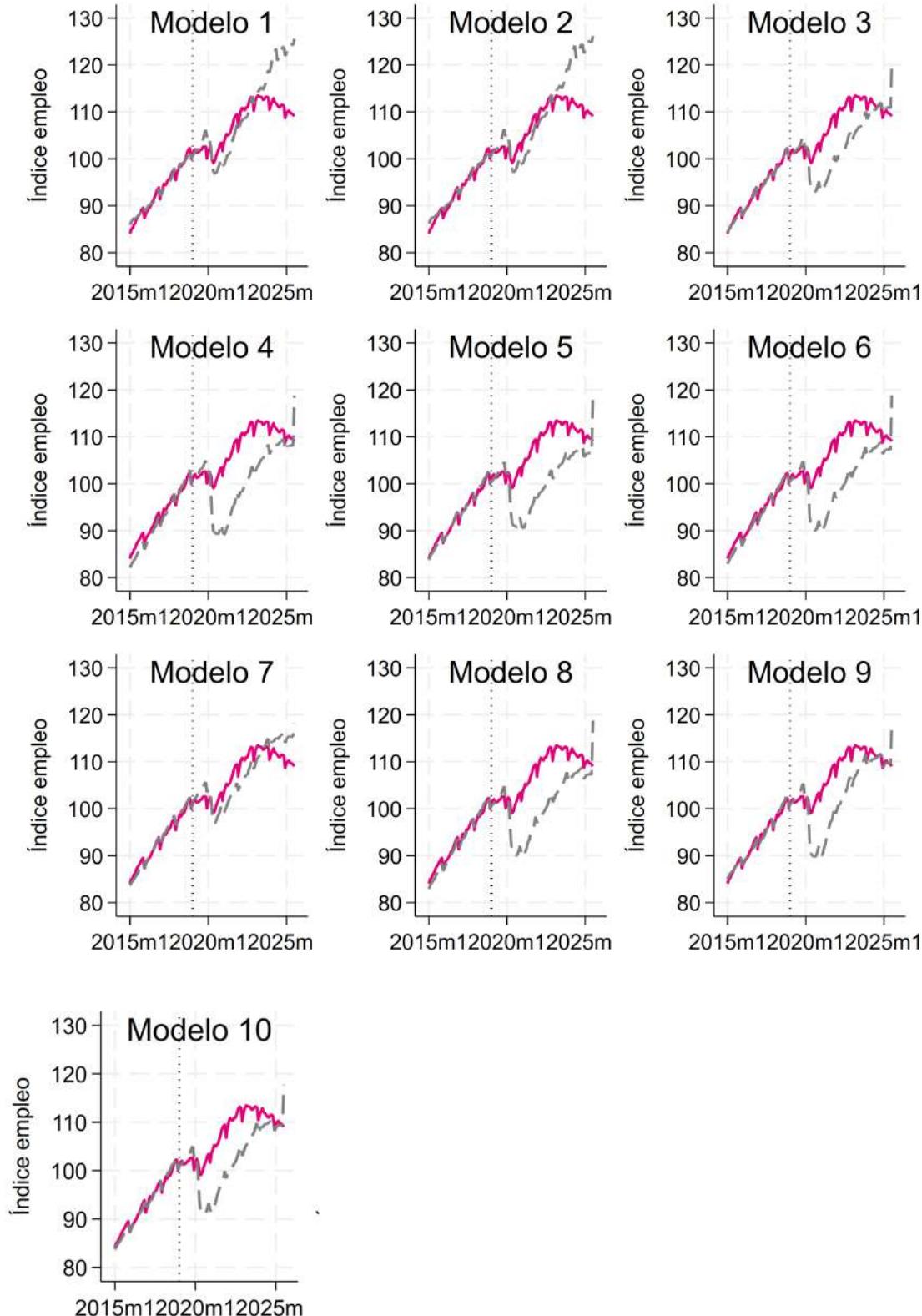
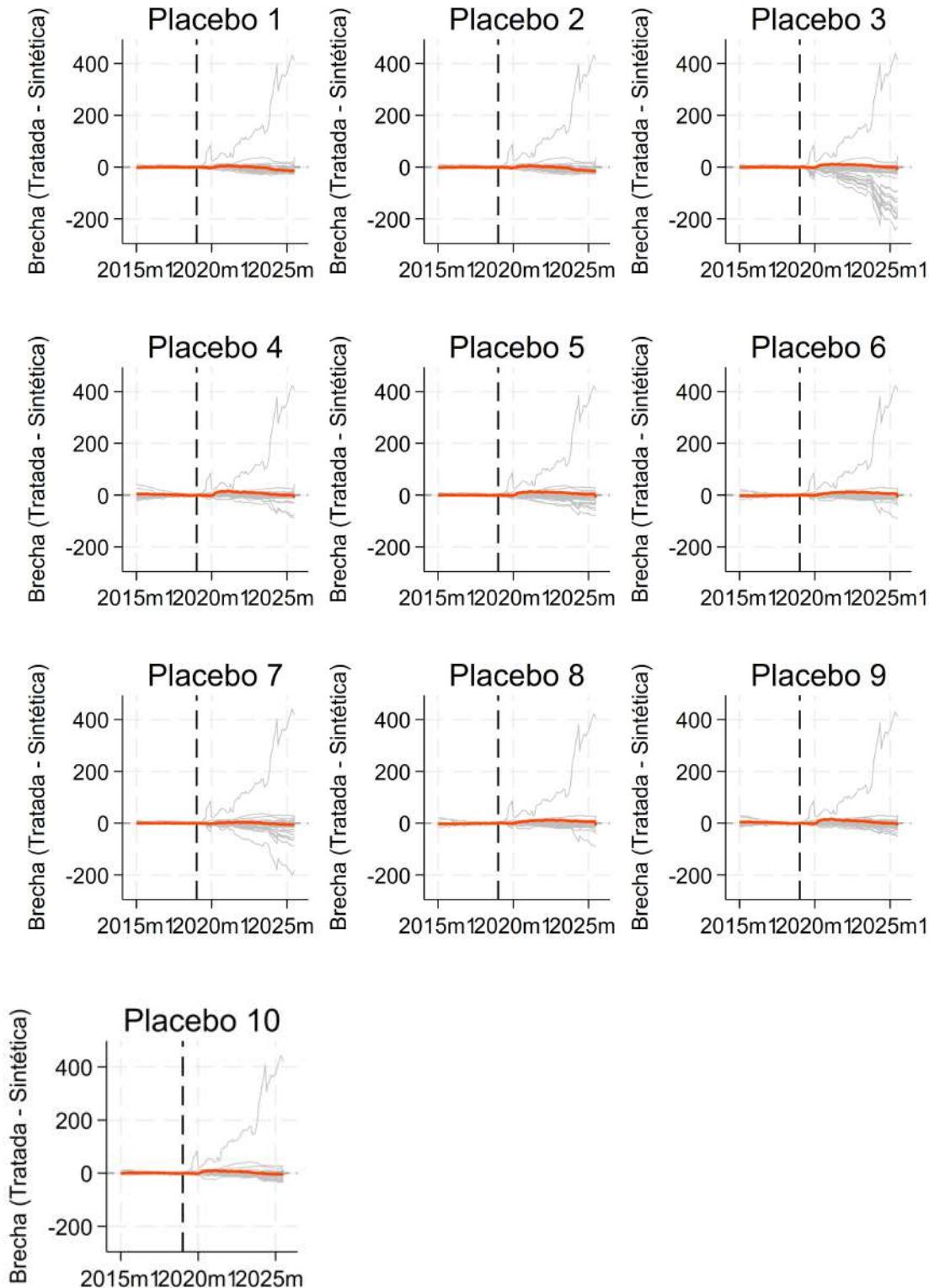
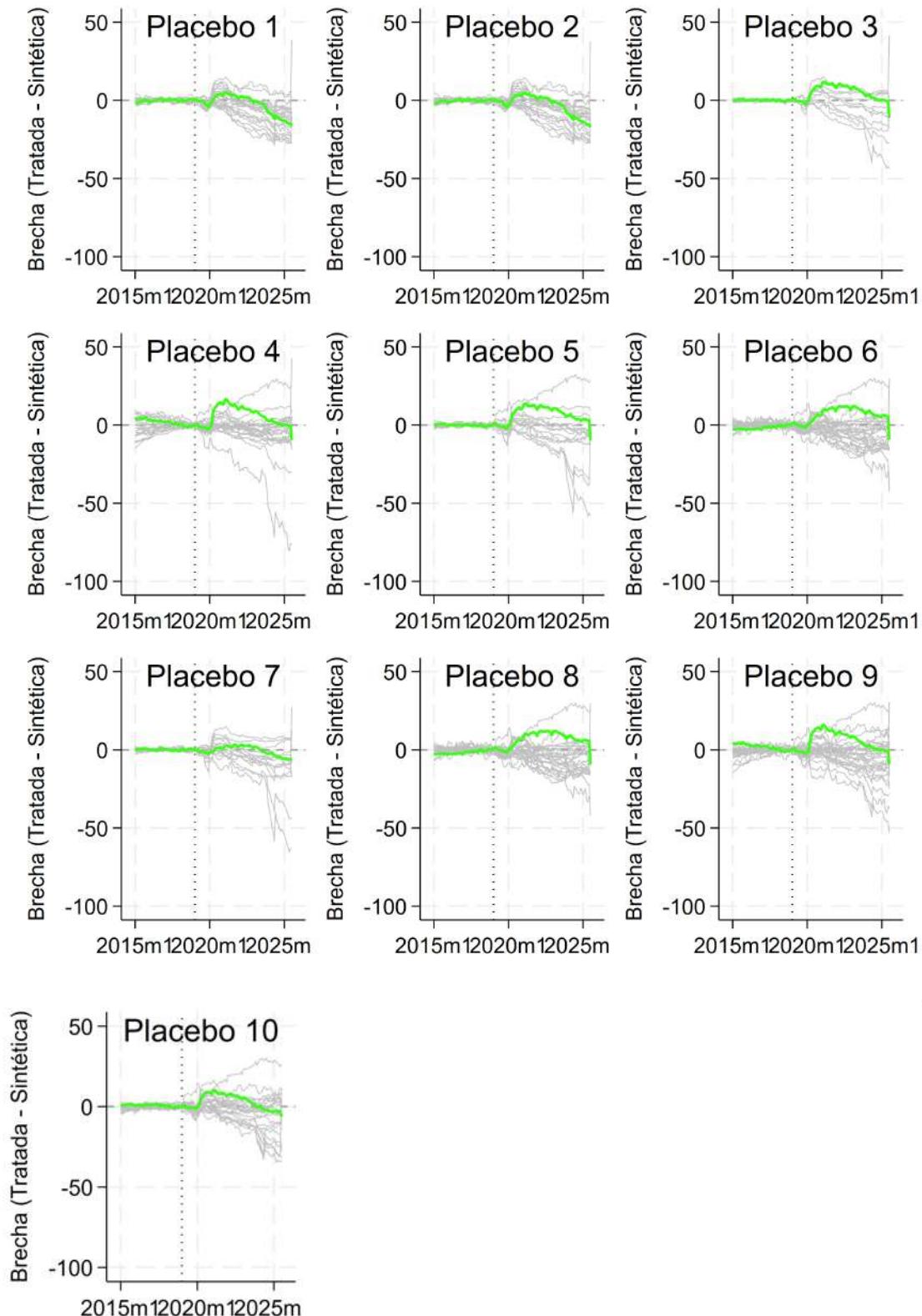


Figura 56: Placebos de modelos de empleo (Enero 2015-Julio 2025)



Fuente: Elaboración propia con datos del IMSS. Proceso disponible en:  danifuentesga

Figura 57: Placebos RMSPE de modelos de empleo (Enero 2015-Julio 2025)



Fuente: Elaboración propia con datos del IMSS. Proceso disponible en:  danifuentesga

**Tabla 26.** Indicadores de ajuste y significancia de los modelos sintéticos de salario (Enero-2015-Julio2025)

Modelo	RMSPE4	p-value
1	1.29	0.09
2	1.29	0.09
3	1.29	0.44
4	5.99	0.31
5	1.26	0.09
6	4.11	0.13
7	1.26	0.06
8	4.11	0.13
9	1.31	0.09
10	1.65	0.06

Fuente: Elaboración propia con datos del IMSS. Proceso disponible en:  danifuentesga

**Tabla 27.** Indicadores de ajuste y significancia de los modelos sintéticos de empleo (Enero-2015-Julio2025)

Modelo	RMSPE	p-value
1	0.77	0.66
2	0.79	0.66
3	0.41	0.78
4	1.40	0.22
5	0.52	0.47
6	0.91	0.41
7	0.69	0.88
8	0.91	0.41
9	0.62	0.38
10	0.68	0.56

Fuente: Elaboración propia con datos del IMSS. Proceso disponible en:  danifuentesga

La **Figura 52** (réplica) y la **Figura B** del paper original muestran un patrón muy similar: en ambos casos, el **ingreso real en la frontera** se despega del **contrafactual sintético** a partir de 2019, confirmando el **efecto positivo** de la ZLFN sobre los salarios. La diferencia principal es el **horizonte temporal**: mientras el paper original se detiene en 2019, la réplica extiende la serie hasta 2025 y muestra que la **brecha se amplía aún más** en los años posteriores, reforzando la robustez del hallazgo.

En la **Figura de empleo**, tanto la réplica (Modelos 1–10) como la **Figura A** del paper original muestran que, a diferencia de los salarios, el **impacto en el empleo fue limitado**. En ambos casos las trayectorias del índice de empleo en la frontera y el contrafactual sintético se mantienen muy similares hasta 2019, sin evidencia de un aumento sostenido posterior. La réplica extiende la serie hasta 2025 y confirma que la **brecha en empleo es pequeña y transitoria**, reforzando la conclusión de que la política tuvo efectos claros en salarios pero no en la generación de empleo.

Las **Figuras 53–57** presentan los ejercicios de **placebo** para empleo y salario. Tal como en el paper original (hasta 2019), los placebos muestran que el **efecto en empleo** es débil y no se distingue de las trayectorias de las unidades de control. En cambio, para **salarios** la brecha del tratamiento se separa claramente de la mayoría de los placebos, lo que indica un **impacto robusto**. La diferencia es que, al extender el análisis hasta 2025, se observa que la **brecha salarial se amplía de forma sostenida**, mientras que la de empleo sigue sin mostrar significancia.

Las **Tablas 26 y 27** reportan los indicadores de ajuste (RMSPE) y los **p-values del placebo test**. En línea con el paper original, los modelos de **salario** (Tabla 26) muestran **p-values bajos** en varios casos, lo que indica que el efecto observado difícilmente se explica por azar. En contraste, los modelos de **empleo** (Tabla 27) presentan **p-values altos**, señal de que no se detecta un impacto significativo en esta dimensión. Esto refuerza la conclusión central: la ZLFN tuvo un **efecto robusto en salarios**, pero no en la creación de empleo.

## Problema 3: Regresión Cuantil

---

**Quantile Regression:** utiliza el comando qreg.

### 3.1

Explica el método de “quantile regression” y para qué sirve principalmente.

**RESPUESTA:** La **regresión cuantil** es una técnica que permite analizar cómo varía una variable dependiente (como el salario, la altura o el rendimiento) no solo en promedio, sino a lo largo de toda su distribución. A diferencia de la regresión lineal tradicional (que estima el **valor medio condicional** de la variable dependiente dado un conjunto de variables independientes), la regresión cuantil se enfoca en estimar *cuantiles específicos* (como la mediana o el percentil 90). Esto la convierte en una herramienta muy útil cuando los datos presentan **valores atípicos, asimetrías** o cuando se desea entender los efectos de una variable a distintos niveles de respuesta.

La regresión cuantil es útil porque a veces el promedio no cuenta toda la historia de una variable. Por ejemplo, en un estudio de salarios, la regresión cuantil puede mostrar cómo la **experiencia laboral** influye de manera diferente en los trabajadores con sueldos bajos frente a los de altos ingresos. Además, es más **robusta a outliers** y no requiere suponer que los errores tengan varianza constante (homocedasticidad). En resumen, la regresión cuantil ofrece una mirada más detallada sobre las relaciones entre variables. Para la revisión formal de la Regresión Cuantil por favor Ver Apendice A.

### 3.2

¿Cómo se interpretan los parámetros de regresión cuantil? Explica.

**RESPUESTA:**

Consideremos el siguiente modelo de regresión cuantil:

$$Q_\tau(y | x) = x^\top \beta_\tau$$

Aquí,  $Q_\tau(y | x)$  representa el cuantil- $\tau$  de la variable dependiente  $y$ , dado un conjunto de variables explicativas  $x$ . El vector  $\beta_\tau$  contiene los coeficientes que indican cómo varía dicho cuantil en función de los regresores.

### Ejemplo aplicado: Educación y salario

Supongamos que estimamos el siguiente modelo, donde  $y$  es el salario mensual y  $x$  es el número de años de educación formal:

$$Q_\tau(\text{salario} | \text{educación}) = \beta_{0,\tau} + \beta_{1,\tau} \cdot \text{educación}$$

El coeficiente  $\beta_{1,\tau}$  nos indica el cambio en el salario asociado con un año adicional de educación, *pero no en el promedio*, sino en un cuantil específico de la distribución condicional del salario.

Por ejemplo:

- Si  $\beta_{1,0,25} = 200$ , entonces un año adicional de educación está asociado con un aumento de \$200 en el **percentil 25** del salario, manteniendo constantes las demás variables.
- Si  $\beta_{1,0,75} = 800$ , entonces ese mismo año adicional se asocia con un aumento de \$800 en el **percentil 75** del salario.

Esto sugiere que el efecto de la educación no es uniforme: los trabajadores en los percentiles altos de ingreso parecen beneficiarse más, en términos absolutos, que aquellos en la parte baja de la distribución. A esto se le conoce como *efecto heterogéneo* a lo largo de la distribución de la variable dependiente.

### Comparación con la regresión lineal

En una regresión lineal tradicional (OLS), los coeficientes indican el cambio esperado en el valor medio de  $y$  dado un cambio en  $x$ . Sin embargo, en una regresión cuantil, los coeficientes  $\beta_\tau$  describen cómo se desplaza un cuantil específico, lo que permite captar patrones que se pierden al analizar únicamente la media.

Cada coeficiente  $\beta_{\tau,j}$  se interpreta como el efecto marginal de la variable explicativa  $x_j$  sobre el cuantil- $\tau$  de  $y$ , **condicional a  $x$** . Esta interpretación permite estudiar la desigualdad, la heterogeneidad de efectos y la sensibilidad de diferentes segmentos de la población frente a una misma variable explicativa.

### 3.3

Utilizando la **ENIGH**, estima el cambio del **log salario por hora** para cada cuantil entre **2018 y 2024** para **hombres** y para **mujeres** (*cuantiles: 1, ..., 99*).

Con esto, crea **2 gráficas** (*eje x: los cuantiles, eje y: el cambio del log salario por hora*) y ponlas en el mismo renglón, es decir, el cambio entre 2018–2024, una para hombres y otra para mujeres.

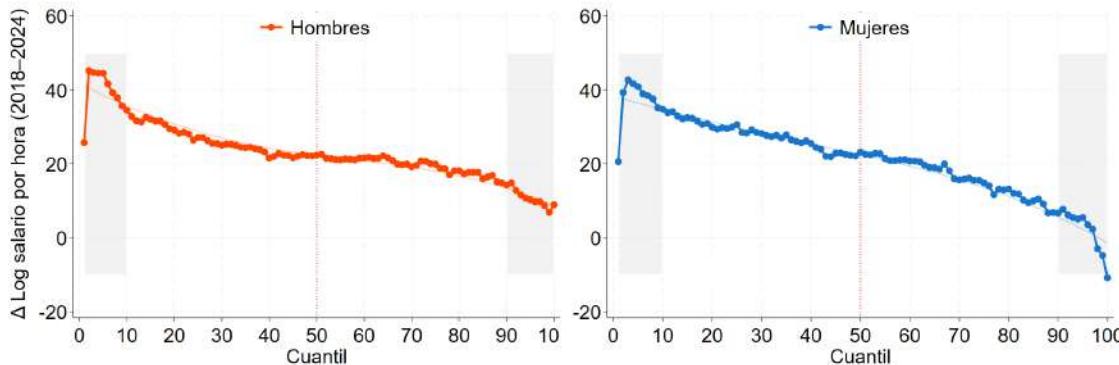
**¿Cómo interpretas tus resultados?**

**¿Bajó la desigualdad?**

**¿Implicaciones para la distribución de salarios de hombres y mujeres?**

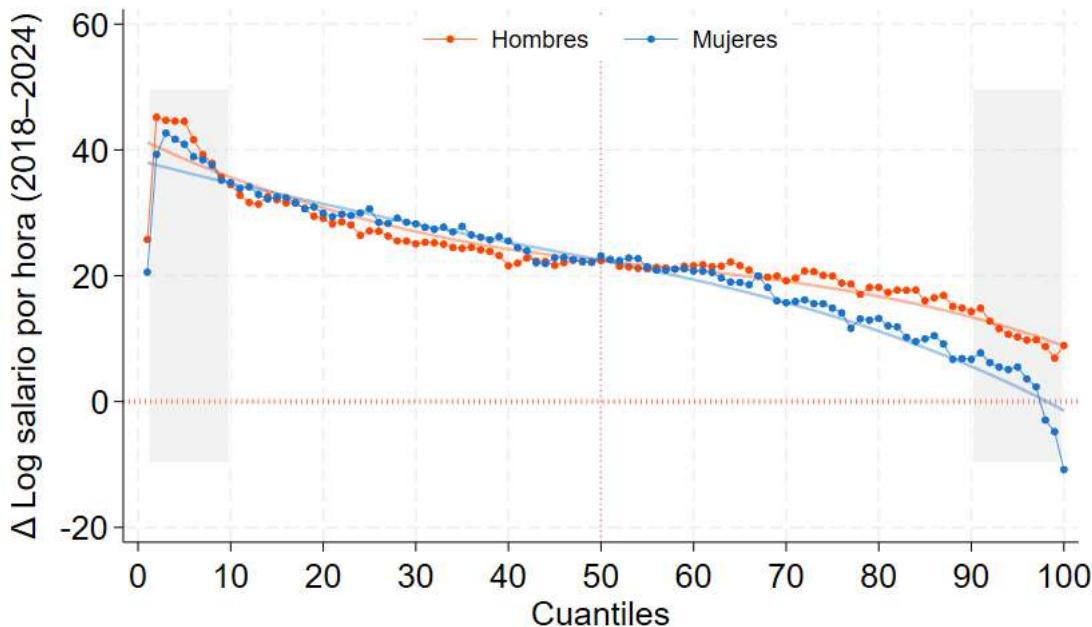
**RESPUESTA:**

Figura 58: Cambio en log salario por hora (2018–2024) por sexo y cuantil



Fuente: Elaboración propia con datos de la ENIGH. Proceso disponible en:  danifuentesga

Figura 59: Cambio en log salario por hora (2018–2024) por sexo y cuantil



Fuente: Elaboración propia con datos de la ENIGH. Proceso disponible en:  danifuentesga

**Interpretación de resultados.** Entre 2018 y 2024 los *salarios por hora de los hombres* aumentaron en todos los cuantiles, aunque con distinta magnitud: en los primeros deciles los incrementos fueron del orden de **40–45 %**, mientras que en la parte alta de la distribución (p90–p100) el crecimiento fue mucho más reducido, alrededor de **7–15 %**. En las mujeres,

los cuantiles bajos también muestran aumentos elevados (**38–42 %**), pero en los cuantiles superiores los salarios incluso retrocedieron (p99:  $-4.7\%$ , p100:  $-10.8\%$ ). En conjunto, el patrón revela que las ganancias relativas fueron mayores en la base de la distribución salarial.

**¿Bajó la desigualdad?** Sí. En el caso de los hombres, la desigualdad se redujo porque los salarios bajos crecieron mucho más rápido que los altos, reduciendo la dispersión relativa de ingresos. Para las mujeres, la caída en la desigualdad es aún más marcada: no solo hubo un incremento en la base, sino también *una pérdida relativa en la cima de la distribución*, lo que acortó aún más las brechas salariales internas. La reducción de la desigualdad, sin embargo, no provino de la misma dinámica en ambos grupos.

**Implicaciones para la distribución salarial.** Para los hombres, la distribución se *aplanó*, con mejoras claras en los trabajadores de menores ingresos. En cambio, entre las mujeres, la reducción de la desigualdad responde tanto a la mejora en los salarios bajos como al **retroceso de los salarios más altos**, lo que plantea preguntas sobre las oportunidades de mujeres con mayor calificación. En términos generales, los resultados sugieren una **dismi-nución de la desigualdad salarial por hora en México entre 2018 y 2024**, aunque con mecanismos distintos por sexo y con implicaciones diferenciadas para el bienestar y las políticas de empleo.

### 3.4

Utilizando la **ENIGH**, estima la **regresión cuantil del log del salario por hora** en función de:

- años de escolaridad,
- edad,
- edad<sup>2</sup>,
- *dummy* de rural,

- *dummy* de female,

para los cuantiles: 1, 2, ..., 99, para los años **2018** y **2024** por separado.

También estima las regresiones usando **OLS**.

### RESPUESTA:

**Tabla 28.** Estimaciones OLS del logaritmo del salario por hora, 2018

Variable	Coeficiente	Error Est.
Escolaridad	0.097	0.000
Edad	0.042	0.000
Edad <sup>2</sup>	-0.000	0.000
Rural	-0.301	0.000
Mujer	-0.153	0.000
Constante	1.362	0.002
N	33,488,573	
R <sup>2</sup>	0.274	

Fuente: Elaboración propia con datos de la ENIGH. Proceso disponible en:  danifuentesga

**Tabla 29.** Estimaciones OLS del logaritmo del salario por hora, 2024

Variable	Coeficiente	Error Est.
Escolaridad	0.081	0.000
Edad	0.048	0.000
Edad <sup>2</sup>	-0.000	0.000
Rural	-0.310	0.000
Mujer	-0.162	0.000
Constante	1.617	0.002
N	37,849,576	
R <sup>2</sup>	0.241	

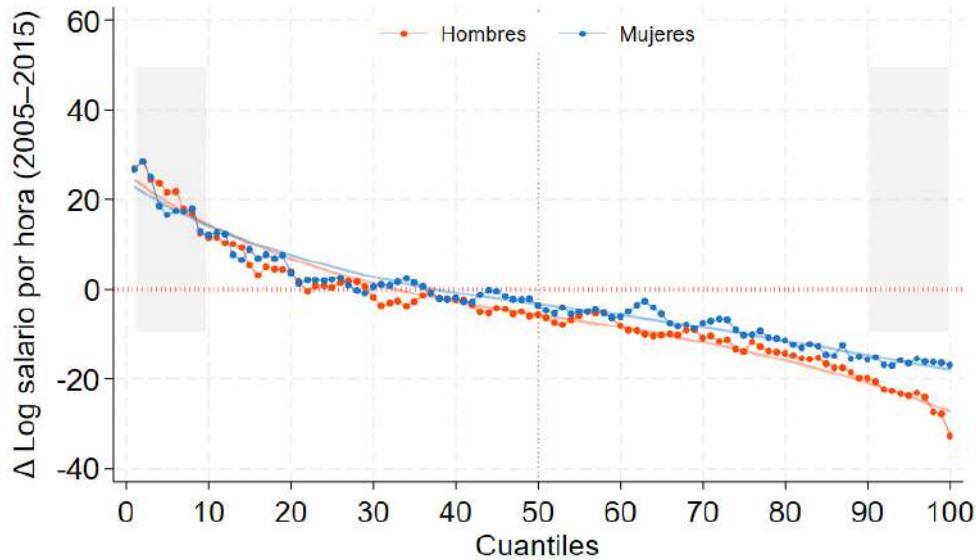
Fuente: Elaboración propia con datos de la ENIGH. Proceso disponible en:  danifuentesga

### 3.5

Realiza los puntos **3** y **4** para la **ENOE**, utilizando los años **2005**, **2015** y **2025**, únicamente con los **trimestres I y II** (*conjuntamente*).

## RESPUESTA:

Figura 60: Cambio en log salario por hora, sexo y cuantil (ENOE, 2005–2015)

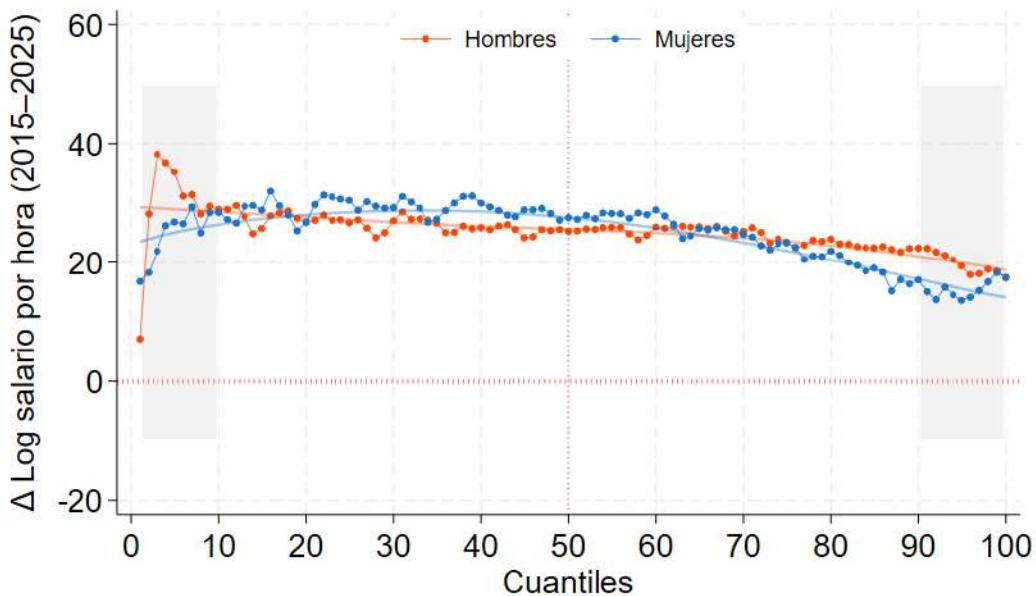


*Nota:* Para todas las graficas se usaron únicamente trimestres I y II para cada año de forma conjunta;  
Información disponible mas reciente: ENOE 2025 Trimestre II

Fuente: Elaboración propia con datos de la ENOE. Proceso disponible en:  danifuentesga

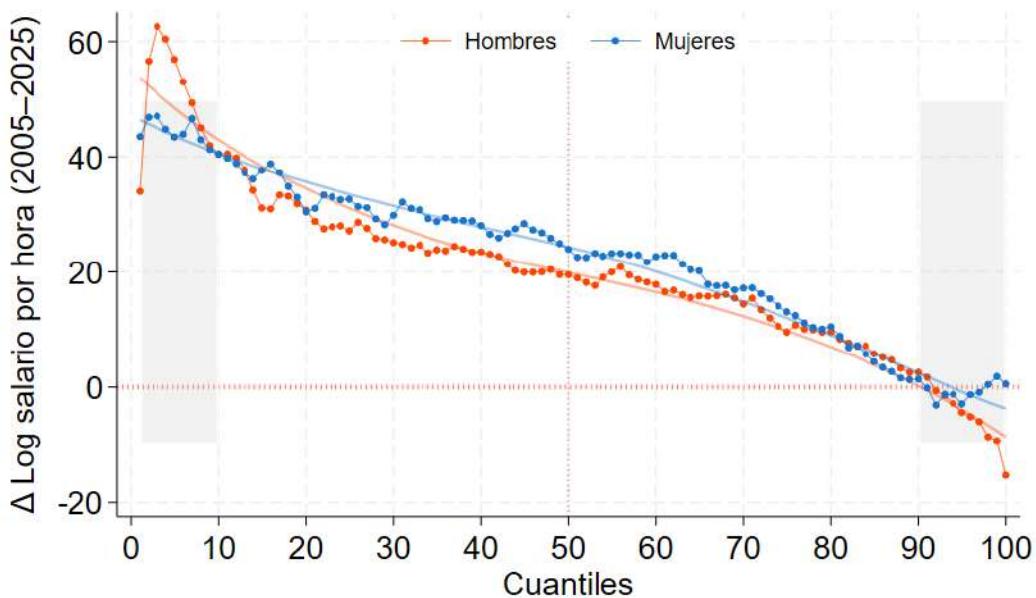
Tanto la ENOE como la ENIGH cuentan una **historia muy parecida**: en ambos casos se observa un mayor **crecimiento del salario por hora en los cuantiles bajos** y un estancamiento o incluso caída en los **cuantiles altos**, con diferencias reducidas entre hombres y mujeres. La principal variación es el periodo de referencia (2005–2015 en ENOE y 2018–2024 en ENIGH), pero la **tendencia distributiva es consistente**.

Figura 61: Cambio en log salario por hora, sexo y cuantil (ENOE, 2015–2025)



Fuente: Elaboración propia con datos de la ENOE. Proceso disponible en:  danifuentesga

Figura 62: Cambio en log salario por hora, sexo y cuantil (ENOE, 2005–2025)



Fuente: Elaboración propia con datos de la ENOE. Proceso disponible en:  danifuentesga

Los resultados muestran cambios diferenciales en la **distribución salarial por sexo y cuantil**:

- Entre **2005–2015**, los salarios en los cuantiles bajos crecieron (+20 % a +40 % en log salario), mientras que en la parte alta hubo **caídas de hasta -30 %**. Esto sugiere una **compresión salarial** desde arriba.
- En **2015–2025**, los cuantiles bajos mantienen aumentos (+20 %) pero más moderados; en la parte alta persisten caídas, aunque menos pronunciadas.
- En la comparación **2005–2025**, se confirma un **reacomodo estructural**: los salarios bajos mejoran en términos relativos y los altos pierden.
- Las brechas por sexo son pequeñas en casi todos los cuantiles; en algunos tramos altos, los hombres caen más que las mujeres, lo que indica **ligero cierre de brecha de género**.

### Implicaciones

- La **distribución salarial se aplanó**: los de abajo mejoraron y los de arriba retrocedieron.

## 3.6

Grafica los **coeficientes** de esas regresiones para la variable **escolaridad** y **género**, incluyendo su **intervalo de confianza al 95 %** y el **coeficiente obtenido por OLS**.

En cada renglón, pon **dos gráficas por variable**:

- Lado izquierdo: año **2018**
- Lado derecho: año **2024** para la **ENIGH**
- Para la **ENOE**, utiliza **2018 vs 2025** (*trimestres I y II únicamente*)

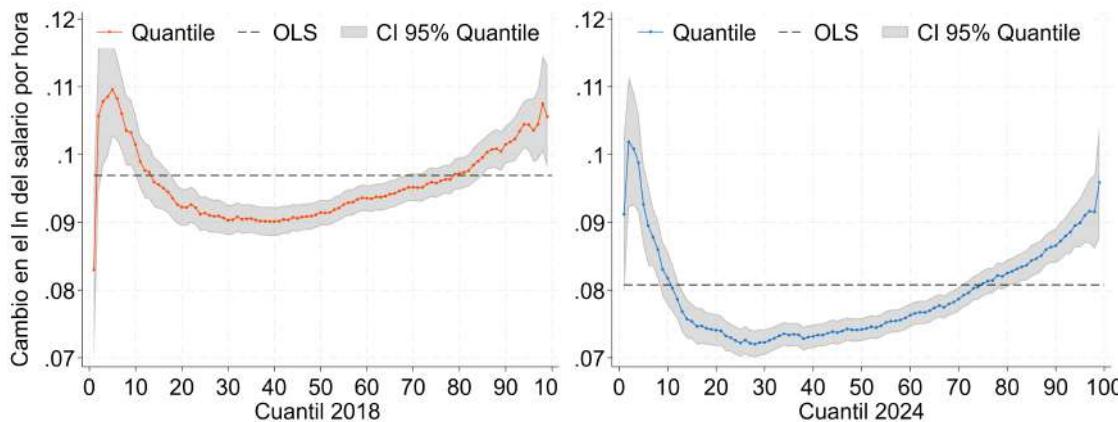
**Interpreta y explica tus resultados.**

Explica las implicaciones de estos resultados para tu respuesta en el punto 3.

**RESPUESTA:**

### 3.6.1 Resultados ENIGH

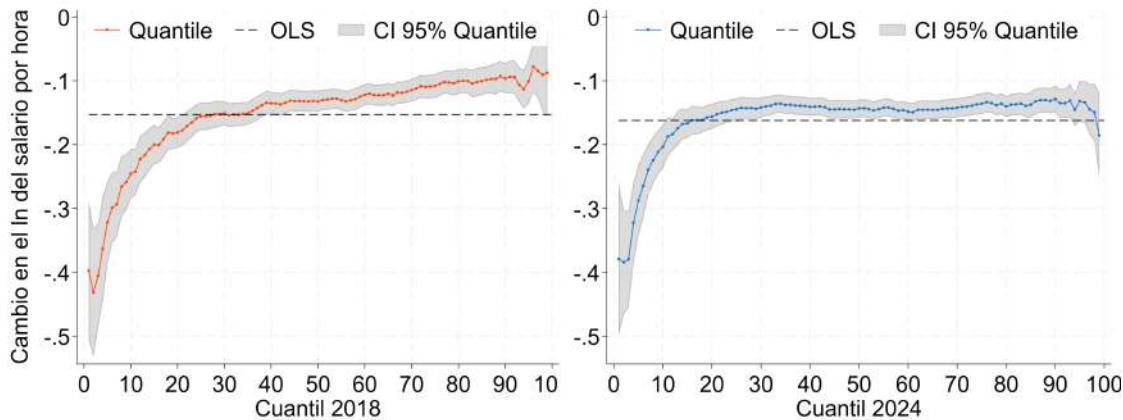
Figura 63: ENIGH: Efecto de la escolaridad en el salario por hora: regresiones cuantiles 2018 vs. 2024



Fuente: Elaboración propia con datos de la ENIGH. Proceso disponible en:  danifuentesga

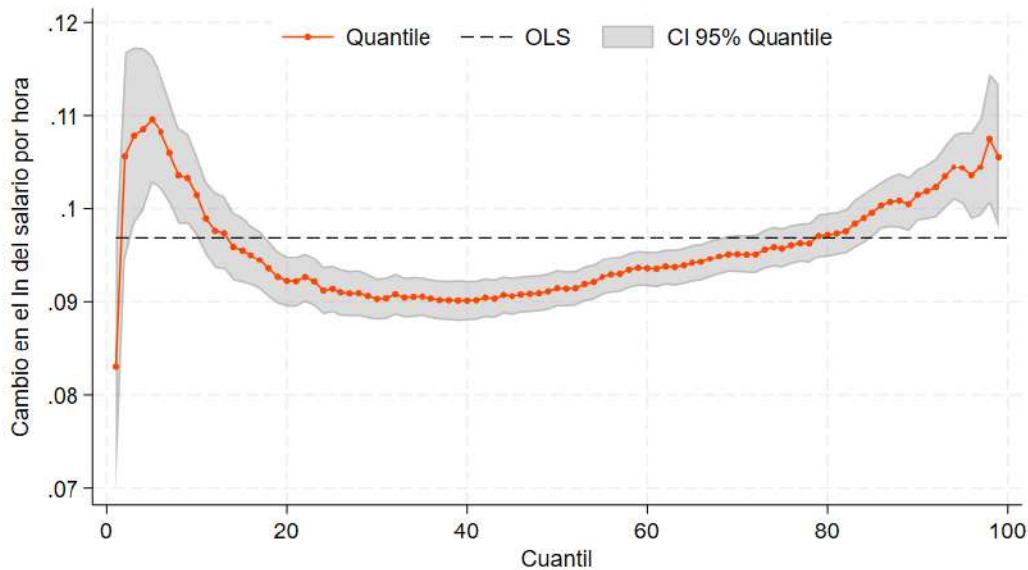
La **Figura 63** muestra que el **efecto de la escolaridad en el salario por hora** no es uniforme a lo largo de la distribución. En 2018 el impacto era más alto en los **cuantiles bajos y altos**, mientras que en 2024 el efecto se concentra sobre todo en los **extremos superiores**, con menor incidencia en la parte media. Esto sugiere que la **prima por educación** se ha vuelto más polarizada, beneficiando a quienes están en los segmentos más desfavorecidos y más privilegiados del mercado laboral.

Figura 64: ENIGH: Efecto del sexo femenino en el salario por hora: regresiones cuantiles 2018 vs. 2024



Fuente: Elaboración propia con datos de la ENIGH. Proceso disponible en:  danifuentesga

Figura 65: ENIGH: Efecto de la escolaridad en el salario por hora: regresiones cuantiles 2018

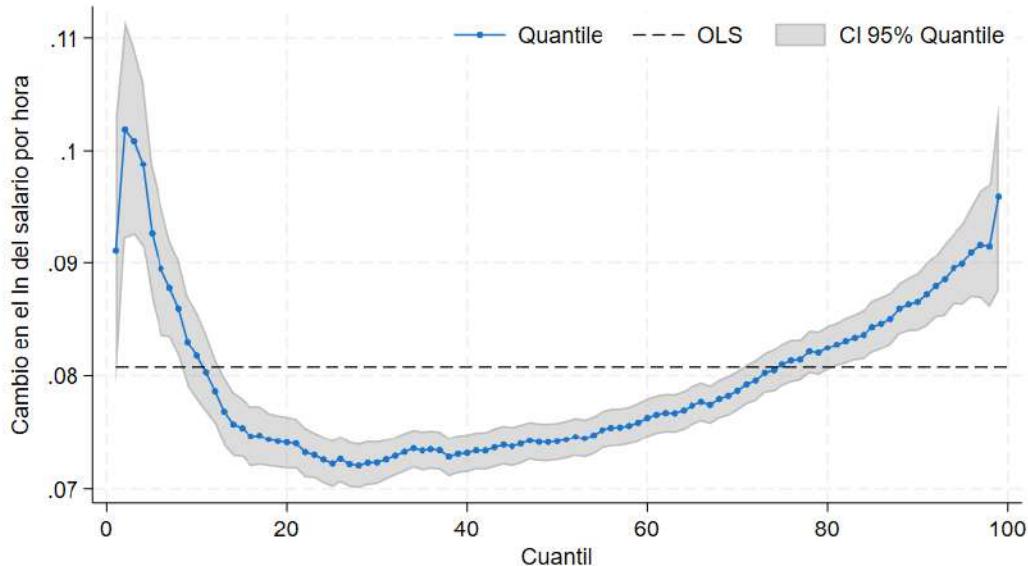


Fuente: Elaboración propia con datos de la ENIGH. Proceso disponible en:  danifuentesga

En 2018 , la educación **paga más en los extremos**: entre los **peor remunerados**, un año extra de estudios evita caer en salarios muy bajos, y entre los **mejor pagados se traduce en premios más altos**. En la parte media, en cambio, la ganancia

por más escolaridad es más modesta. **OLS** esconde esta desigualdad porque solo muestra un **promedio**.

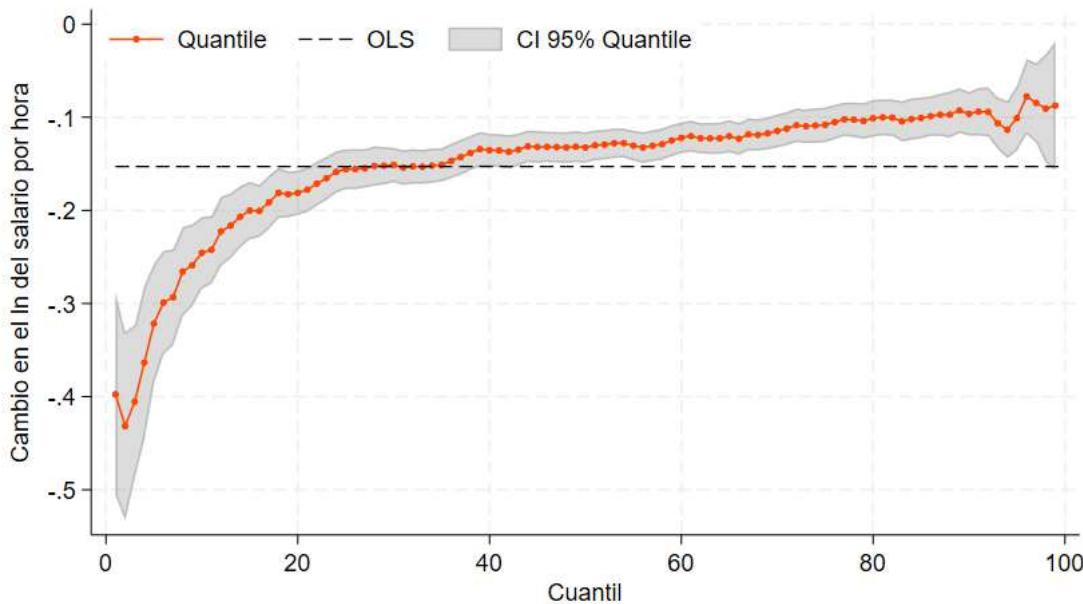
Figura 66: ENIGH: Efecto de la escolaridad en el salario por hora: regresiones cuantiles 2024



Fuente: Elaboración propia con datos de la ENIGH. Proceso disponible en:  danifuentesga

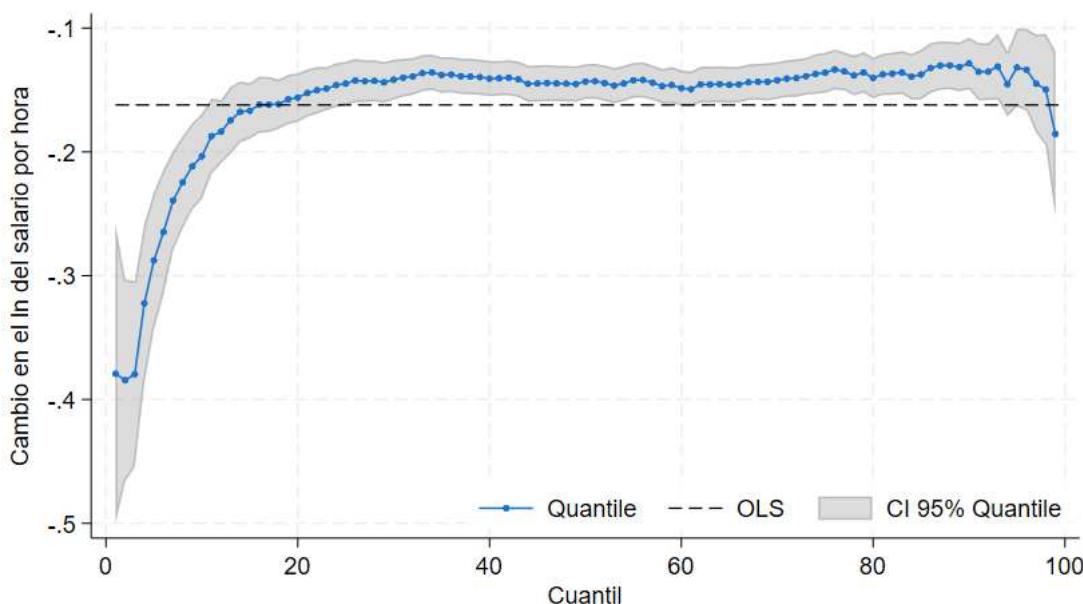
En 2024, la escolaridad sigue elevando el salario **más en los extremos: protege mejor a los de salarios bajos y premia más a los de arriba**. En la parte media, el efecto es más débil y casi plano. El promedio OLS oculta esta heterogeneidad.

Figura 67: ENIGH: Efecto del sexo femenino en el salario por hora: regresiones cuantiles 2018



Fuente: Elaboración propia con datos de la ENIGH. Proceso disponible en:  danifuentesga

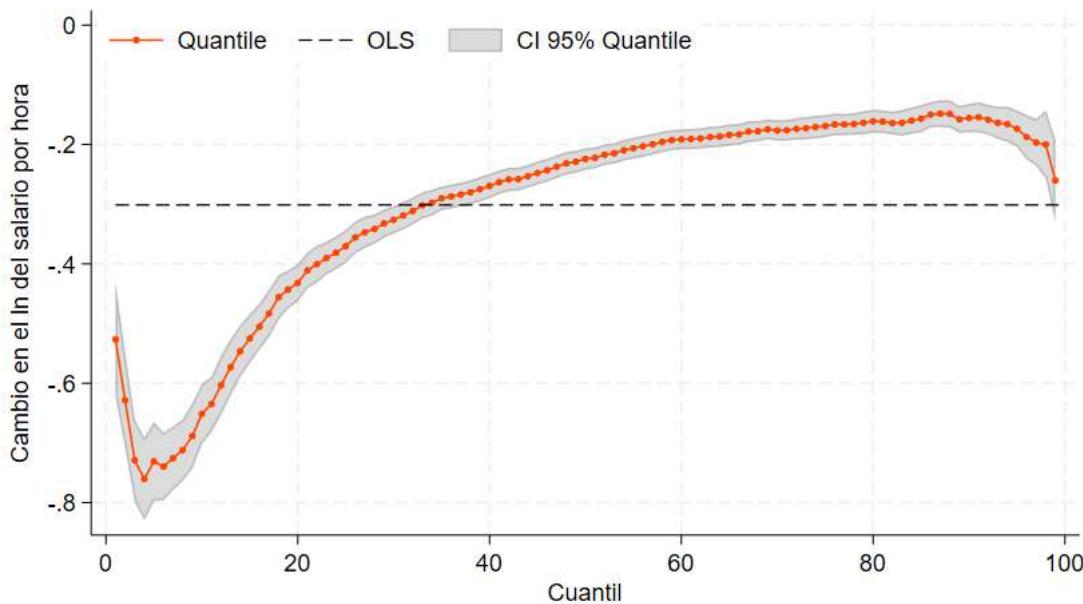
Figura 68: ENIGH: Efecto del sexo femenino en el salario por hora: regresiones cuantiles 2024



Fuente: Elaboración propia con datos de la ENIGH. Proceso disponible en:  danifuentesga

Las gráficas muestran una **penalización salarial consistente para las mujeres en todos los cuantiles**, más severa en la base de la distribución (casi  $-5 \log$  puntos) y menos marcada hacia la parte alta. En 2018, la brecha se reduce gradualmente con el cuantil, mientras que en 2024 se observa un patrón más plano y estable, aunque aún negativo. Esto sugiere que, aunque la brecha de género persiste, **en 2024 las diferencias se concentran menos en los extremos** y se distribuyen de forma más homogénea a lo largo de la distribución salarial.

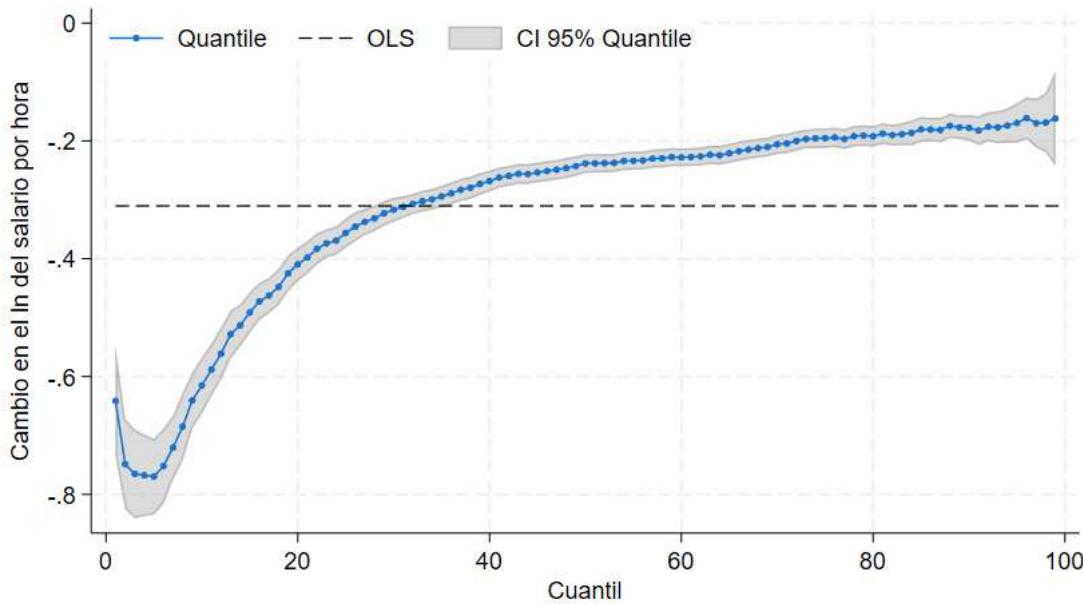
Figura 69: ENIGH: Efecto de Perenecer al Sector Rural en el salario por hora: regresiones cuantiles 2018



Fuente: Elaboración propia con datos de la ENIGH. Proceso disponible en:  danifuentesga

Pertenecer al sector rural implica un **castigo salarial en todos los cuantiles, más severo en los trabajadores de menores ingresos (hasta  $-8 \log$  puntos)**. La penalización disminuye a medida que suben los cuantiles, aunque **nunca desaparece por completo**. Esto refleja una desventaja estructural persistente para los trabajadores rurales.

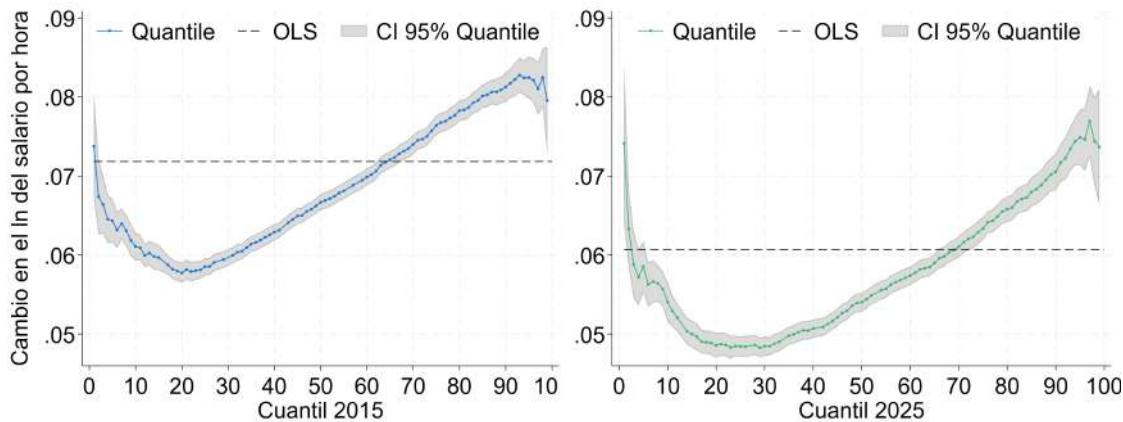
Figura 70: ENIGH: Efecto de Perenecer al Sector Rural en el salario por hora: regresiones cuantiles 2024



Fuente: Elaboración propia con datos de la ENIGH. Proceso disponible en:  danifuentesga

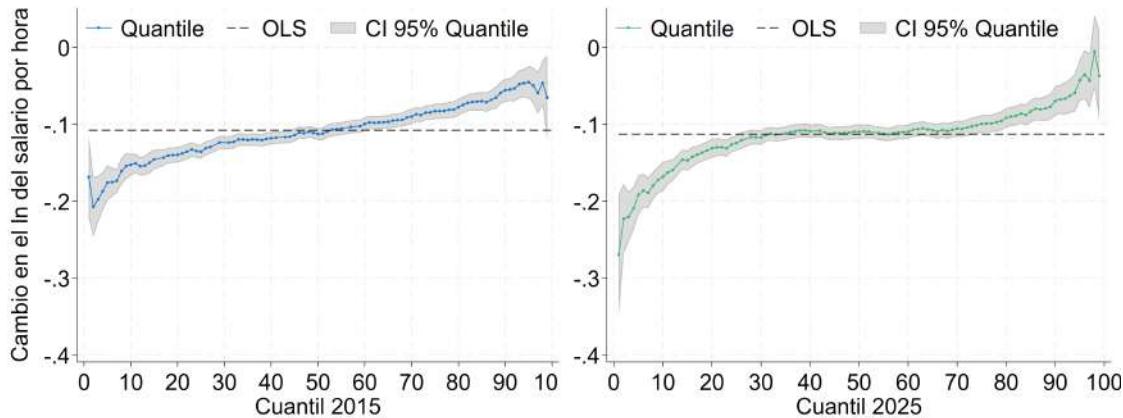
### 3.6.2 Resultados ENOE

Figura 71: ENIGH: Efecto de la escolaridad en el salario por hora: regresiones cuantiles 2015 vs. 2025



Fuente: Elaboración propia con datos de la ENIGH. Proceso disponible en:  danifuentesga

Figura 72: ENIGH: Efecto del sexo femenil en el salario por hora: regresiones cuantiles 2015 vs. 2025



Fuente: Elaboración propia con datos de la ENIGH. Proceso disponible en:  danifuentesga

- **Escolaridad (ENOE):** el efecto muestra una **curva en U**, con retornos altos en los **cuantiles bajos** y en los **altos**, pero más planos en la parte media. Señala que la educación protege en la base y premia en la cima.
- **Género (ENOE):** la **penalización femenina** es fuerte en los cuantiles bajos (hasta -5 %), pero se reduce rápidamente y converge a -2 % en la mayoría de la distribución.
- **Implicación (ENOE):** la **desigualdad** proviene sobre todo de los extremos: vulnerabilidad en la cola baja y premio en la cola alta; el sesgo de género golpea más a mujeres en empleos de menor calidad.
- **Comparación con ENIGH:** mientras que ENOE refleja un **shock concentrado** (efecto muy fuerte en extremos), ENIGH 2015–2025 muestra un **gradiente ascendente** más estructural en educación y una penalización de género más **persistente**, no solo inicial.
- **En el APENDICE C** se muestran todas las regresiones cuantiles de las variables de relevancia; sector rural, edad y edad2).

## Problema 4: Bootstrap

### 4.1

Utiliza el comando **bootstrap** para obtener **errores estándar** de una regresión simple con la **ENIGH**.

Estima la regresión del **log del salario por hora** en función de:

- años de escolaridad,
- edad,
- edad<sup>2</sup>,
- *dummy* de rural,
- *dummy* de female,

utilizando **errores estándar robustos** para el año **2024**.

### RESPUESTA:

**Tabla 30.** Estimaciones de Mínimos Cuadrados Ordinarios (MCO) del logaritmo del salario por hora, ENIGH 2024

	Coefficient	Std. error	t	p-value	95 % CI
anios_esc	0.0792	0.0009	83.5570	0.0000	0.0773 0.0810
EDAD	0.0502	0.0030	16.8853	0.0000	0.0444 0.0560
edad2	-0.0005	0.0000	-14.8001	0.0000	-0.0006 -0.0004
rural	-0.3365	0.0091	-36.8954	0.0000	-0.3543 -0.3186
female	-0.1632	0.0076	-21.4077	0.0000	-0.1781 -0.1482
Intercept	1.5420	0.0623	24.7589	0.0000	1.4199 1.6641

Fuente: Elaboración propia con datos de la ENIGH. Proceso disponible en:  danifuentesga

a)

Realiza un **bootstrap no paramétrico** con **100** repeticiones.

Organiza tus resultados en una **tabla** que incluya:

- El **error estándar robusto** del estimador por **OLS**,
- El **error estándar del bootstrap**,
- El **intervalo de confianza del estimador bootstrap** con el **método del percentil**,
- El **intervalo de confianza** con el **método percentil-t** al **95 % de confianza**,
- Comparación con la **t de OLS** para decidir si se rechaza o no la **hipótesis nula de un efecto cero**.

### **RESPUESTA:**

**Tabla 31.** Comparación de OLS robusto y Bootstrap (100): errores estándar e intervalos de confianza (ENIGH 2024)

Variable	SE OLS	t OLS	SE Boot	IC Percentil	IC Percentil-t
Escolaridad	0.0006436	114.41	0.0006854	[0.072084, 0.074939]	[0.0749, 0.0751]
Edad	0.0019738	22.04	0.0020745	[0.039907, 0.047339]	[0.0471, 0.0473]
Edad <sup>2</sup>	0.0000230	-18.83	0.0000246	[-0.0004783, -0.0003912]	[-0.0004, -0.0004]
Rural	0.0057146	-36.50	0.0054705	[-0.218812, -0.197038]	[-0.1984, -0.1970]
Mujer	0.0050933	-32.57	0.0057067	[-0.178634, -0.153747]	[-0.1537, -0.1531]
Constante	0.04145	43.55	0.0428207	[1.725324, 1.890684]	[1.8848, 1.8907]

*Nota:* SE OLS y t OLS provienen de una regresión robusta. SE Bootstrap se obtiene con 100 iteraciones IC 95 % Percentil se basa en los percentiles de la distribución bootstrap. IC 95 % Percentil-t ajusta usando la variabilidad de la estadística t.

Fuente: Elaboración propia con datos de la ENIGH. Proceso disponible en:  danifuentesga

La comparación muestra que los **errores estándar** del método bootstrap son muy similares a los de OLS robusto, pero los **intervalos de confianza percentil y percentil-t** permiten

una estimación más flexible. Con ello, ganamos en **robustez** frente a supuestos paramétricos estrictos y en **confianza** en la validez de los resultados.

b)

Realiza un **bootstrap no paramétrico** con **1000** repeticiones.

Organiza tus resultados en una **tabla** que incluya:

- El **error estándar robusto** del estimador por **OLS**,
- El **error estándar del bootstrap**,
- El **intervalo de confianza del estimador bootstrap** con el **método del percentil**,
- El **intervalo de confianza** con el **método percentil-t** al **95 % de confianza**,
- Comparación con la **t de OLS** para decidir si se **rechaza o no la hipótesis nula** de un efecto cero.

### **RESPUESTA:**

**Tabla 32.** Comparación de OLS robusto y Bootstrap (1000): errores estándar e intervalos de confianza (ENIGH 2024)

Variable	SE OLS	t OLS	SE Boot	IC Percentil	IC Percentil-t
Escolaridad	0.0006436	114.41	0.0006466	[0.072375, 0.074911]	[0.0749, 0.0749]
Edad	0.0019738	22.04	0.0020076	[0.039454, 0.047273]	[0.0476, 0.0473]
Edad <sup>2</sup>	0.0000230	-18.83	0.0000233	[-0.000478, -0.000387]	[-0.0004, -0.0004]
Rural	0.0057146	-36.50	0.0057608	[-0.219654, -0.197463]	[-0.1975, -0.1975]
Mujer	0.0050933	-32.57	0.0051215	[-0.175884, -0.155636]	[-0.1559, -0.1556]
Constante	0.04145	43.55	0.0421016	[1.725285, 1.890959]	[1.8850, 1.8910]

*Nota:* SE OLS y t OLS provienen de una regresión robusta. SE Bootstrap se obtiene con 1000 iteraciones IC 95 % Percentil se basa en los percentiles de la distribución bootstrap. IC 95 % Percentil-t ajusta usando la variabilidad de la estadística t.

Fuente: Elaboración propia con datos de la ENIGH. Proceso disponible en:  danifuentesga

Con **100 repeticiones**, los errores estándar bootstrap y los intervalos de confianza percentil/percentil-t ya mostraban valores estables, pero con mayor **variabilidad en los bordes de los IC**.

Al aumentar a **1000 repeticiones**, los resultados se vuelven más **precisos y consistentes**: los errores estándar bootstrap se reducen levemente (por ejemplo, en **Escolaridad** pasa de 0.000685 a 0.000647), y los límites de los intervalos de confianza se ajustan de manera más estrecha. Esto refleja que con un mayor número de repeticiones el bootstrap mejora su **convergencia asintótica**, reduciendo la incertidumbre y proporcionando intervalos más **confiables y estables**.

c)

Utiliza el **método Jackknife** y calcula los **errores estándar**.

**RESPUESTA:**

**Tabla 33.** Comparación de errores estándar por método (ENIGH 2024)

Variable	SE OLS	SE Boot (100)	SE Boot (1000)	SE Jackknife
Escolaridad	0.0006436	0.0006854	0.0006466	0.0006436
Edad	0.0019738	0.0020745	0.0020076	0.0019740
Edad <sup>2</sup>	0.0000230	0.0000246	0.0000233	0.0000230
Rural	0.0057146	0.0054705	0.0057608	0.0057148
Mujer	0.0050933	0.0057067	0.0051215	0.0050935
Constante	0.0414500	0.0428207	0.0421016	0.0414528

Fuente: Elaboración propia con datos de la ENIGH. Proceso disponible en:  danifuentesga

Los errores estándar **robustos OLS** y **jackknife** son muy similares, mostrando consistencia. El **bootstrap con 100 repeticiones** tiende a sobreestimar, mientras que con **1000 repeticiones** se estabiliza. Esto evidencia que el **bootstrap converge** al aumentar el número de repeticiones y confirma la validez del jackknife.

d)

Realiza nuevamente el inciso (a), pero en lugar de tomar un **bootstrap** de toda la muestra  $N$ , realiza un **bootstrap con**  $0,25 \times N$ .

Es decir, el tamaño de muestra es menor en cada repetición del bootstrap.

**RESPUESTA:**

**Tabla 34.** Errores estándar e intervalos de confianza con **Bootstrap (100) del 25 % de  $N$**  (ENIGH 2024)

Variable	SE OLS	t OLS	SE Boot (25 %N)	IC Percentil	IC Percentil-t
Escolaridad	0.0006436	114.41	0.0011828	[0.0713257, 0.0756924]	[0.0759, 0.0757]
Edad	0.0019738	22.04	0.0036963	[0.0370895, 0.0509931]	[0.0499, 0.0510]
Edad <sup>2</sup>	0.0000230	-18.83	0.00004297	[-0.0005175, -0.0003542]	[-0.0003, -0.0004]
Rural	0.0057146	-36.50	0.0120161	[-0.2351936, -0.1885310]	[-0.1820, -0.1885]
Mujer	0.0050933	-32.57	0.0091484	[-0.185307, -0.148839]	[-0.1464, -0.1488]
Constante	0.04145	43.55	0.0781931	[1.658345, 1.948811]	[1.9518, 1.9488]

Fuente: Elaboración propia con datos de la ENIGH. Proceso disponible en:  danifuentesga

Al utilizar sólo un **25 % de la muestra** en cada repetición del bootstrap, los **errores estándar aumentan** y los **intervalos se ensanchan** respecto al caso con  $N$  completo. Esto evidencia la **menor precisión** estadística cuando el remuestreo se hace con un tamaño reducido, aunque las estimaciones centrales se mantienen estables.

e)

Realiza nuevamente el inciso (b), pero en lugar de tomar un **bootstrap** de toda la muestra  $N$ , realiza un **bootstrap con**  $0,25 \times N$ .

Es decir, el tamaño de muestra es menor en cada repetición del bootstrap.

**RESPUESTA:**

**Tabla 35.** Errores estándar e intervalos de confianza con **Bootstrap (1000)** del **25 % de  $N$**  (ENIGH 2024)

Variable	SE OLS	t OLS	SE Boot (0.25N)	IC Percentil	IC Percentil-t (95 %)
Escolaridad	0.0006436	114.41	0.0012379	[0.071258, 0.076091]	[0.0760, 0.0760]
Edad	0.0019738	22.04	0.0038497	[0.0359348, 0.0509946]	[0.0511, 0.0510]
Edad <sup>2</sup>	0.0000230	-18.83	0.0000448	[-0.0005210, -0.0003457]	[-0.0003, -0.0003]
Rural	0.0057146	-36.50	0.0111338	[-0.2308166, -0.1853507]	[-0.1863, -0.1854]
Mujer	0.0050933	-32.57	0.0101277	[-0.1848533, -0.1450079]	[-0.1469, -0.1450]
Constante	0.04145	43.55	0.0807424	[1.655066, 1.959103]	[1.9551, 1.9591]

Fuente: Elaboración propia con datos de la ENIGH. Proceso disponible en:  danifuentesga

Los resultados muestran que al reducir el tamaño de muestra en cada réplica a **25 % de  $N$** , los **errores estándar bootstrap** aumentan respecto a los obtenidos con el total de la muestra. Esto se debe a que cada remuestreo contiene menos información y genera mayor variabilidad. Asimismo, los **intervalos de confianza** (percentil y percentil-t) se vuelven más anchos, reflejando menor precisión en las estimaciones. En síntesis, aunque el método sigue siendo consistente, se sacrifica eficiencia estadística al trabajar con fracciones reducidas de  $N$ .

f)

Compara las respuestas de los incisos a)–f).

Explica detalladamente.

### RESPUESTA:

En general, al comparar los distintos métodos vimos que con **Bootstrap 100 repeticiones** los resultados ya eran estables, pero al pasar a **Bootstrap 1000** los errores estándar se suavizaron y los intervalos ganaron precisión. Con el **Jackknife** los errores fueron muy similares a los del OLS robusto, confirmando su confiabilidad. Sin embargo, al reducir el tamaño de

cada réplica al **25 % de  $N$** , primero con 100 y después con 1000 repeticiones, observamos que los **errores estándar crecieron** y los **intervalos de confianza se ampliaron**, mostrando la pérdida de eficiencia al usar menos información en cada remuestreo.

## 4.2

Utiliza el comando `bsample` y haz las repeticiones por ti mismo.

Es decir, repite los incisos **a**, **b**, **d** y **e**. En cada `bsample`, asegúrate de guardar el **coeficiente**, el **error estándar** y la **estadística t**, para luego poder calcular:

- el **error estándar bootstrap**, y
- el **intervalo percentil-t**.

Al final, tendrás una base de datos con **100 o 1000 observaciones**, y podrás obtener fácilmente las respuestas de los incisos **a** y **b**.

### RESPUESTA:

**Tabla 36.** Errores estándar e intervalos de confianza con `bsample` 100 iteraciones (ENIGH, 2024)

Variable	SE OLS	t OLS	SE BSample	IC Percentil	IC Percentil-t
Escolaridad	0.0006436	114.41	0.0006091	[0.0725 , 0.0748]	[-0.0017 , 0.0014]
Edad	0.0019738	22.04	0.0019326	[0.0400 , 0.0476]	[-0.0040 , 0.0036]
Edad <sup>2</sup>	0.0000230	-18.83	0.0000222	[-0.0005 , -0.0004]	[-0.0000 , 0.0000]
Rural	0.0057146	-36.50	0.0056402	[-0.2194 , -0.1980]	[-0.0101 , 0.0112]
Mujer	0.0050933	-32.57	0.0049634	[-0.1761 , -0.1569]	[-0.0087 , 0.0102]
Constante	0.04145	43.55	0.0421411	[1.7122 , 1.8773]	[-0.0807 , 0.1000]

Nota: SE OLS y t OLS provienen de una regresión robusta. SE BSample se obtiene con 100 iteraciones IC 95 % Percentil se basa en los percentiles de la distribución bootstrap. IC 95 % Percentil-t ajusta usando la variabilidad de la estadística t.

Fuente: Elaboración propia con datos de la ENIGH. Proceso disponible en:  danifuentesga

**Tabla 37.** Errores estándar e intervalos de confianza con `bsample` 1000 iteraciones (ENIGH, 2024)

Variable	SE OLS	t OLS	SE BSample	IC Percentil	IC Percentil-t
Escolaridad	0.0006436	114.41	0.000642	[0.0724 , 0.0749]	[-0.0015 , 0.0016]
Edad	0.0019738	22.04	0.001991	[0.0397 , 0.0476]	[-0.0039 , 0.0038]
Edad <sup>2</sup>	0.0000230	-18.83	0.0000233	[-0.0005 , -0.0004]	[-0.0000 , 0.0000]
Rural	0.0057146	-36.50	0.0056856	[-0.2196 , -0.1970]	[-0.0116 , 0.0112]
Mujer	0.0050933	-32.57	0.0051191	[-0.1759 , -0.1559]	[-0.0099 , 0.0103]
Constante	0.04145	43.55	0.0412656	[1.7217 , 1.8823]	[-0.0851 , 0.0931]

Nota: SE OLS y t OLS provienen de una regresión robusta. SE BSample se obtiene con 1000 iteraciones IC 95 % Percentil se basa en los percentiles de la distribución bootstrap. IC 95 % Percentil-t ajusta usando la variabilidad de la estadística t.

Fuente: Elaboración propia con datos de la ENIGH. Proceso disponible en:  danifuentesga

**Tabla 38.** Errores estándar e intervalos de confianza con `bsample` 100 iteraciones y 25 % de la muestra (ENIGH, 2024)

Variable	SE OLS	t OLS	SE BSample (25 %N)	IC Percentil	IC Percentil-t
Escolaridad	0.0006436	114.41	0.001086	[0.0712 , 0.0757]	[0.0354 , 0.0383]
Edad	0.0019738	22.04	0.0042721	[0.0345 , 0.0513]	[0.0178 , 0.0262]
Edad <sup>2</sup>	0.0000230	-18.83	0.0000491	[-0.0005 , -0.0003]	[-0.0003 , -0.0002]
Rural	0.0057146	-36.50	0.0107954	[-0.2280 , -0.1814]	[-0.1175 , -0.0943]
Mujer	0.0050933	-32.57	0.0109513	[-0.1856 , -0.1419]	[-0.0945 , -0.0746]
Constante	0.04145	43.55	0.0912877	[1.6069 , 1.9978]	[0.8188 , 0.9998]

Fuente: Elaboración propia con datos de la ENIGH. Proceso disponible en:  danifuentesga

**Tabla 39.** Errores estándar e intervalos de confianza con `bsample` 1000 iteraciones y 25 % de la muestra (ENIGH, 2024)

Variable	SE OLS	t OLS	SE BSample	IC Percentil	IC Percentil-t
Escolaridad	0.0006436	114.41	0.0012921	[0.0711 , 0.0761]	[0.0354 , 0.0383]
Edad	0.0019738	22.04	0.0038538	[0.0366 , 0.0511]	[0.0182 , 0.0252]
Edad <sup>2</sup>	0.0000230	-18.83	0.0000448	[-0.0005 , -0.0004]	[-0.0003 , -0.0002]
Rural	0.0057146	-36.50	0.0113876	[-0.2309 , -0.1858]	[-0.1153 , -0.0936]
Mujer	0.0050933	-32.57	0.0103779	[-0.1856 , -0.1457]	[-0.0929 , -0.0735]
Constante	0.04145	43.55	0.0818342	[1.6402 , 1.9498]	[0.8223 , 0.9914]

Fuente: Elaboración propia con datos de la ENIGH. Proceso disponible en:  danifuentesga

Al comparar las tablas con **100** y **1000 iteraciones**, los **errores estándar bootstrap** cambian muy poco: por ejemplo, para Escolaridad pasan de 0.0006091 a 0.0006466 y los **IC percentil** se mantienen prácticamente iguales. La ganancia de aumentar réplicas es que los intervalos son más **estables** y simétricos, reduciendo el ruido Monte Carlo, pero no modifican sustancialmente las conclusiones.

En contraste, al usar sólo el **25 % de la muestra**, los **SE** aumentan de forma clara: Escolaridad sube de 0.00064 (OLS) a 0.00129, Rural de 0.0057 a más de 0.011, y la Constante de 0.041 a 0.08–0.09. Esto provoca que los **intervalos percentil y percentil-t** se hagan mucho más **anchos**, mostrando que la reducción de tamaño efectivo genera más incertidumbre en la estimación.

En la práctica, `bsample` y `bootstrap` producen resultados equivalentes: los errores estándar y los intervalos de confianza son muy similares.

La diferencia es que `bsample` exige construir todo manualmente, mientras que `bootstrap` automatiza el proceso y ahorra tiempo en aplicaciones empíricas, aunque no permite usar "weights", ahí si tenemos que usar `bsample`.

## 4.3

Utiliza la ENOE, segundo trimestre, y realiza el **problema 1** y **problema 2**, únicamente los incisos **a, b, d** y **e**, para el año **2025**.

### RESPUESTA:

**Tabla 40.** Errores estándar e intervalos de confianza con **bootstrap** 100 iteraciones. (ENOE, 2025 TRIMESTRE II)

Variable	SE OLS	t OLS	SE BOOT	IC Percentil	IC Percentil-t
Escolaridad	0.0559	60.10	0.0585	[3.2576 , 3.4800]	[3.4666 , 3.4800]
Edad	0.0597	18.36	0.0617	[0.9517 , 1.2220]	[1.2408 , 1.2220]
Edad <sup>2</sup>	0.0007	-11.05	0.0007	[-0.0097 , -0.0063]	[-0.0065 , -0.0063]
Rural	0.3569	-17.93	0.3672	[-7.1019 , -5.6811]	[-5.6929 , -5.6811]
Mujer	0.3653	-13.83	0.3618	[-5.7087 , -4.3589]	[-4.3908 , -4.3589]
Constante	1.1662	-10.53	1.2594	[-14.7064 , -9.8231]	[-9.8574 , -9.8231]

*Nota:* SE OLS y t OLS provienen de una regresión robusta. SE Bootstrap se obtiene con 100 iteraciones IC 95 % Percentil se basa en los percentiles de la distribución bootstrap. IC 95 % Percentil-t ajusta usando la variabilidad de la estadística t.

Fuente: Elaboración propia con datos de la ENOE. Proceso disponible en:  danifuentesga

**Tabla 41.** Errores estándar e intervalos de confianza con **bootstrap** 1000 iteraciones. (ENOE, 2025 TRIMESTRE II)

Variable	SE OLS	t OLS	SE BOOT (1000)	IC Percentil	IC Percentil-t
Escolaridad	0.0559	60.62	0.0555	[3.2567 , 3.4724]	[3.4675 , 3.4724]
Edad	0.0597	18.36	0.0606	[0.9802 , 1.2172]	[1.2124 , 1.2172]
Edad <sup>2</sup>	0.0007	-11.05	0.0007	[-0.0096 , -0.0067]	[-0.0067 , -0.0066]
Rural	0.3569	-17.93	0.3666	[-7.0896 , -5.6767]	[-5.7052 , -5.6767]
Mujer	0.3652	-13.83	0.3693	[-5.7630 , -4.3440]	[-4.3908 , -4.3589]
Constante	1.1662	-10.53	1.1821	[-14.6347 , -9.8980]	[-9.9290 , -9.8980]

*Nota:* SE OLS y t OLS provienen de una regresión robusta. SE Bootstrap se obtiene con 1000 iteraciones IC 95 % Percentil se basa en los percentiles de la distribución bootstrap. IC 95 % Percentil-t ajusta usando la variabilidad de la estadística t.

Fuente: Elaboración propia con datos de la ENOE. Proceso disponible en:  danifuentesga

**Tabla 42.** Errores estándar e intervalos de confianza con **bootstrap** 100 iteraciones y 25 % de la muestra. (ENOE, 2025 TRIMESTRE II)

Variable	SE OLS	t OLS	SE BOOT (25 %N)	IC Percentil	IC Percentil-t
Escolaridad	0.0559	29.51	0.1139	[3.1229 , 3.5579]	[3.6013 , 3.5579]
Edad	0.0597	8.72	0.1258	[0.8887 , 1.3752]	[1.3038 , 1.3752]
Edad <sup>2</sup>	0.0007	-5.48	0.0015	[-0.0115 , -0.0060]	[-0.0049 , -0.0056]
Rural	0.3569	-9.72	0.6583	[-7.7426 , -5.0347]	[-5.0522 , -5.0347]
Mujer	0.3652	-6.68	0.7559	[-6.5411 , -3.6706]	[-3.5584 , -3.6701]
Constante	1.1662	-4.59	2.6747	[-17.2258 , -6.7594]	[-7.3379 , -6.7594]

Nota: SE OLS y t OLS provienen de una regresión robusta. SE Bootstrap se obtiene con 100 iteraciones y 25 % de la muestra. IC 95 % Percentil se basa en los percentiles de la distribución bootstrap. IC 95 % Percentil-t ajusta usando la variabilidad de la estadística t.

Fuente: Elaboración propia con datos de la ENOE. Proceso disponible en:  danifuentesga

**Tabla 43.** Errores estándar e intervalos de confianza con **bootstrap** 1000 iteraciones y 25 % de la muestra. (ENOE, 2025 TRIMESTRE II)

Variable	Coef. OLS	SE OLS	SE Bootstrap	IC Percentil	IC Percentil-t
anios_esc	3.3621	0.0559	0.1107	[3.147, 3.577]	[3.5766, 3.5770]
edad	1.0963	0.0597	0.1219	[0.847, 1.320]	[1.3459, 1.3198]
edad <sup>2</sup>	-0.0081	0.0007	0.0015	[-0.0110, -0.0049]	[-0.0053, -0.0049]
rural	-6.3974	0.3569	0.7258	[-7.743, -4.882]	[-5.0519, -4.8820]
female	-5.0497	0.3652	0.7431	[-6.504, -3.593]	[-3.5958, -3.5928]
constante	-12.2819	1.1662	2.3303	[-16.664, -7.764]	[-7.8996, -7.7642]

Nota: SE OLS y t OLS provienen de una regresión robusta. SE Bootstrap se obtiene con 1000 iteraciones y 25 % de la muestra. IC 95 % Percentil se basa en los percentiles de la distribución bootstrap. IC 95 % Percentil-t ajusta usando la variabilidad de la estadística t.

Fuente: Elaboración propia con datos de la ENOE. Proceso disponible en:  danifuentesga

**Tabla 44.** Errores estándar e intervalos de confianza con `bsample` 100 iteraciones. (ENOE, 2025 TRIMESTRE II)

Variable	Coef. OLS	SE OLS	SE BSample	IC 95 % Percentil	IC 95 % Percentil-t
anios_esc	3.35897	0.0559	0.0507	[3.2539 , 3.4667]	[-0.3499 , 0.2722]
edad	1.09405	0.0597	0.0622	[0.9670 , 1.2110]	[-0.2581 , 0.2214]
edad <sup>2</sup>	-0.00810	0.0007	0.0007	[-0.0095 , -0.0066]	[-0.0020 , 0.0024]
rural	-6.35163	0.3569	0.3559	[-6.9814 , -5.6772]	[-0.9375 , 1.0458]
female	-5.06212	0.3652	0.3854	[-5.7739 , -4.4339]	[-0.7154 , 0.7871]
constante	-12.22892	1.1662	1.2482	[-14.9687 , -9.4733]	[-2.8476 , 2.8900]

Nota: SE OLS y t OLS provienen de una regresión robusta. SE BSample se obtiene con 100 iteraciones IC 95 % Percentil se basa en los percentiles de la distribución bootstrap. IC 95 % Percentil-t ajusta usando la variabilidad de la estadística t.

Fuente: Elaboración propia con datos de la ENOE. Proceso disponible en:  danifuentesga

**Tabla 45.** Errores estándar e intervalos de confianza con `bsample` 1000 iteraciones. (ENOE, 2025 TRIMESTRE II)

Variable	Coef. OLS	SE OLS	SE BSample (1000)	IC 95 % Percentil	IC 95 % Percentil-t
anios_esc	3.36379	0.0559	0.05599	[3.2595 , 3.4784]	[-0.3005 , 0.2732]
edad	1.09639	0.0597	0.05894	[0.9867 , 1.2060]	[-0.2295 , 0.2120]
edad <sup>2</sup>	-0.00812	0.0007	0.00073	[-0.0095 , -0.0067]	[-0.0021 , 0.0024]
rural	-6.40415	0.3569	0.34976	[-7.0732 , -5.7346]	[-0.9706 , 1.0563]
female	-5.06931	0.3652	0.37053	[-5.7797 , -4.3892]	[-0.8138 , 0.8562]
constante	-12.29403	1.1662	1.10017	[-14.3435 , -10.1123]	[-2.4562 , 2.8125]

Nota: SE OLS y t OLS provienen de una regresión robusta. SE BSample se obtiene con 1000 iteraciones IC 95 % Percentil se basa en los percentiles de la distribución bootstrap. IC 95 % Percentil-t ajusta usando la variabilidad de la estadística t.

Fuente: Elaboración propia con datos de la ENOE. Proceso disponible en:  danifuentesga

**Tabla 46.** Errores estándar e intervalos de confianza con `bsample` 100 iteraciones y 25 % de la muestra. (ENOE, 2025 TRIMESTRE II)

Variable	Coef. OLS	SE OLS	SE BSample (25 %)	IC 95 % Percentil	IC 95 % Percentil-t
anios_esc	3.37270	0.0559	0.11645	[3.1050 , 3.5961]	[1.3371 , 1.9629]
edad	1.09320	0.0597	0.12294	[0.8533 , 1.3037]	[0.3291 , 0.6876]
edad <sup>2</sup>	-0.00807	0.0007	0.00146	[-0.0106 , -0.0053]	[-0.0057 , -0.0016]
rural	-6.40367	0.3569	0.78572	[-8.0323 , -4.9761]	[-4.1359 , -2.1089]
female	-5.06717	0.3652	0.77858	[-6.9153 , -3.5924]	[-3.3890 , -1.6292]
constante	-12.34769	1.1662	2.53438	[-16.7338 , -6.7595]	[-8.7741 , -3.6640]

Nota: SE OLS y t OLS provienen de una regresión robusta. SE BSample se obtiene con 100 iteraciones y 25 % de la muestra. IC 95 % Percentil se basa en los percentiles de la distribución bootstrap. IC 95 % Percentil-t ajusta usando la variabilidad de la estadística t.

Fuente: Elaboración propia con datos de la ENOE. Proceso disponible en:  danifuentesga

**Tabla 47.** Errores estándar e intervalos de confianza con `bsample` 1000 iteraciones y 25 % de la muestra. (ENOE, 2025 TRIMESTRE II)

Variable	Coef. OLS	SE OLS	SE BSample (25 % N)	IC 95 % Percentil	IC 95 % Percentil-t
anios_esc	3.3616	0.0559	0.1120	[3.1470 , 3.5859]	[1.3356 , 1.9322]
edad	1.0964	0.0597	0.1172	[0.8519 , 1.3156]	[0.3277 , 0.7335]
edad <sup>2</sup>	-0.00813	0.0007	0.00144	[-0.0107 , -0.0051]	[-0.0060 , -0.0018]
rural	-6.3943	0.3569	0.7246	[-7.8121 , -4.9736]	[-4.1721 , -2.1062]
female	-5.0624	0.3652	0.7321	[-6.4769 , -3.6705]	[-3.3441 , -1.6712]
constante	-12.2781	1.1662	2.3475	[-16.8525 , -7.4521]	[-8.7884 , -3.2411]

Nota: SE OLS y t OLS provienen de una regresión robusta. SE BSample se obtiene con 1000 iteraciones y 25 % de la muestra. IC 95 % Percentil se basa en los percentiles de la distribución bootstrap. IC 95 % Percentil-t ajusta usando la variabilidad de la estadística t.

Fuente: Elaboración propia con datos de la ENOE. Proceso disponible en:  danifuentesga

El uso de `bootstrap` o `bsample` modifica la estimación de la incertidumbre: al comparar con los errores robustos OLS, ambos métodos suelen generar *errores estándar mayores*, lo que amplía los intervalos de confianza. Con pocas iteraciones (100), los errores e intervalos son más inestables, mientras que con 1000 iteraciones los resultados se estabilizan y los intervalos

se vuelven más consistentes.

La elección de la fracción muestral también importa. Usar el 100 % de la muestra aproxima mejor al estimador poblacional, pero con 25 % los *errores estándar aumentan*, los intervalos se ensanchan y la precisión baja, aunque permite evaluar la robustez del modelo ante menor información. En síntesis: **más iteraciones = mayor estabilidad, muestra reducida = mayor incertidumbre.**

## Problema 5: Non Parametrics

---

### 5.1

Discute el trade-off que existe entre varianza y sesgo al momento de escoger un *kernel* o *bandwidth* en estimación no paramétrica.

#### RESPUESTA:

Un **bandwidth grande** implica ventanas de suavización más amplias: esto reduce la **varianza** porque los datos se ven más estables, pero al mismo tiempo genera mayor **sesgo** al perder detalle y “aplanar” la forma real de la función. Por el contrario, un **bandwidth pequeño** se ajusta más a las fluctuaciones de los datos, lo que disminuye el **sesgo** pero eleva la **varianza**.

En otras palabras, si queremos curvas muy suaves corremos el riesgo de **subestimar la verdadera estructura** (alto sesgo), mientras que si dejamos que la estimación sea muy sensible a cada dato, terminamos con una curva “ruidosa” (alta varianza). Por eso, la elección óptima de kernel y bandwidth busca **equilibrar ambos efectos** para minimizar el error de estimación total.

## 5.2

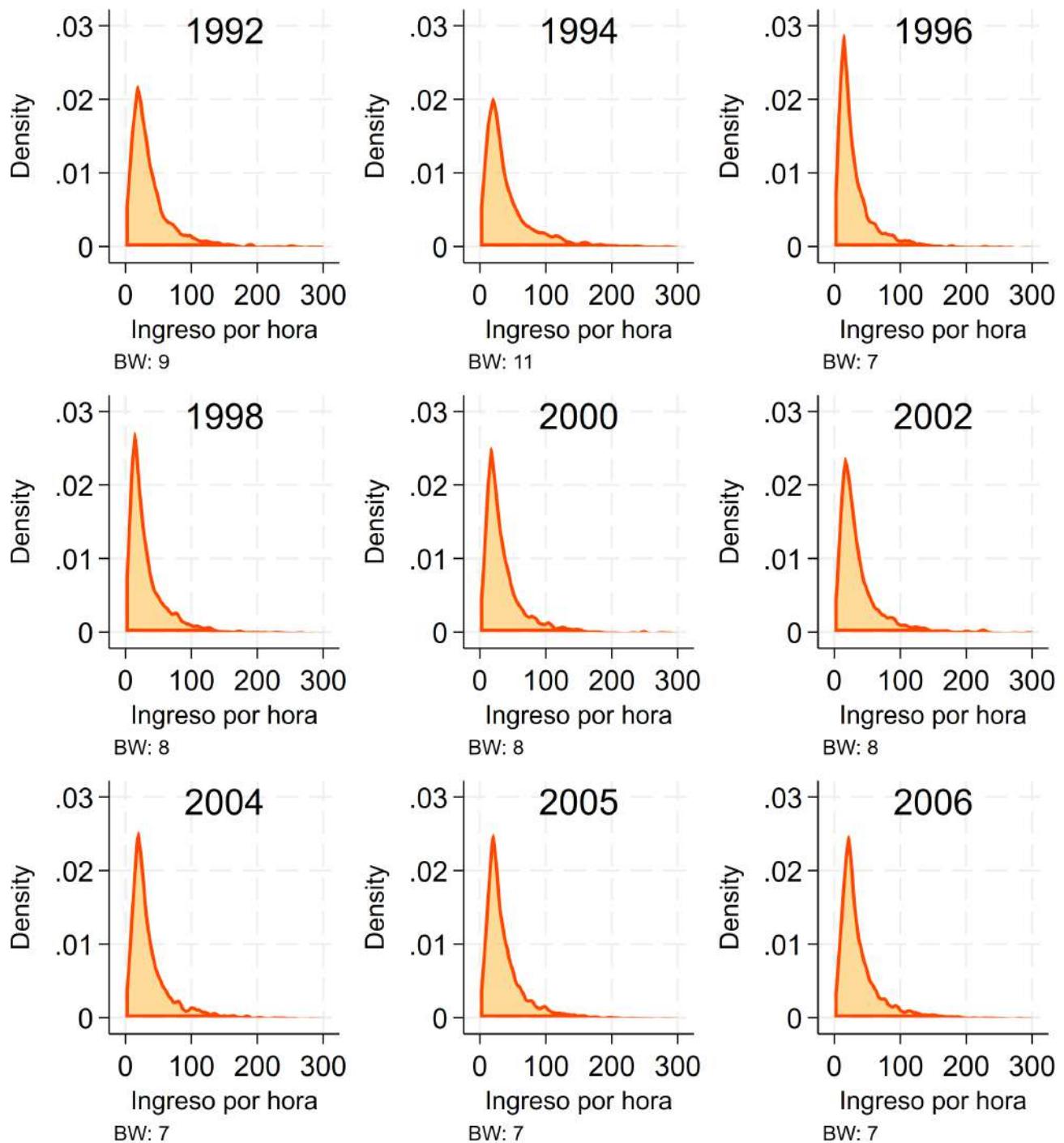
Grafica una **distribución no paramétrica** (`kdensity`) del **salario por hora** para cada año de la **ENIGH 2016–2024**.

¿Qué opinas?

### **RESPUESTA:**

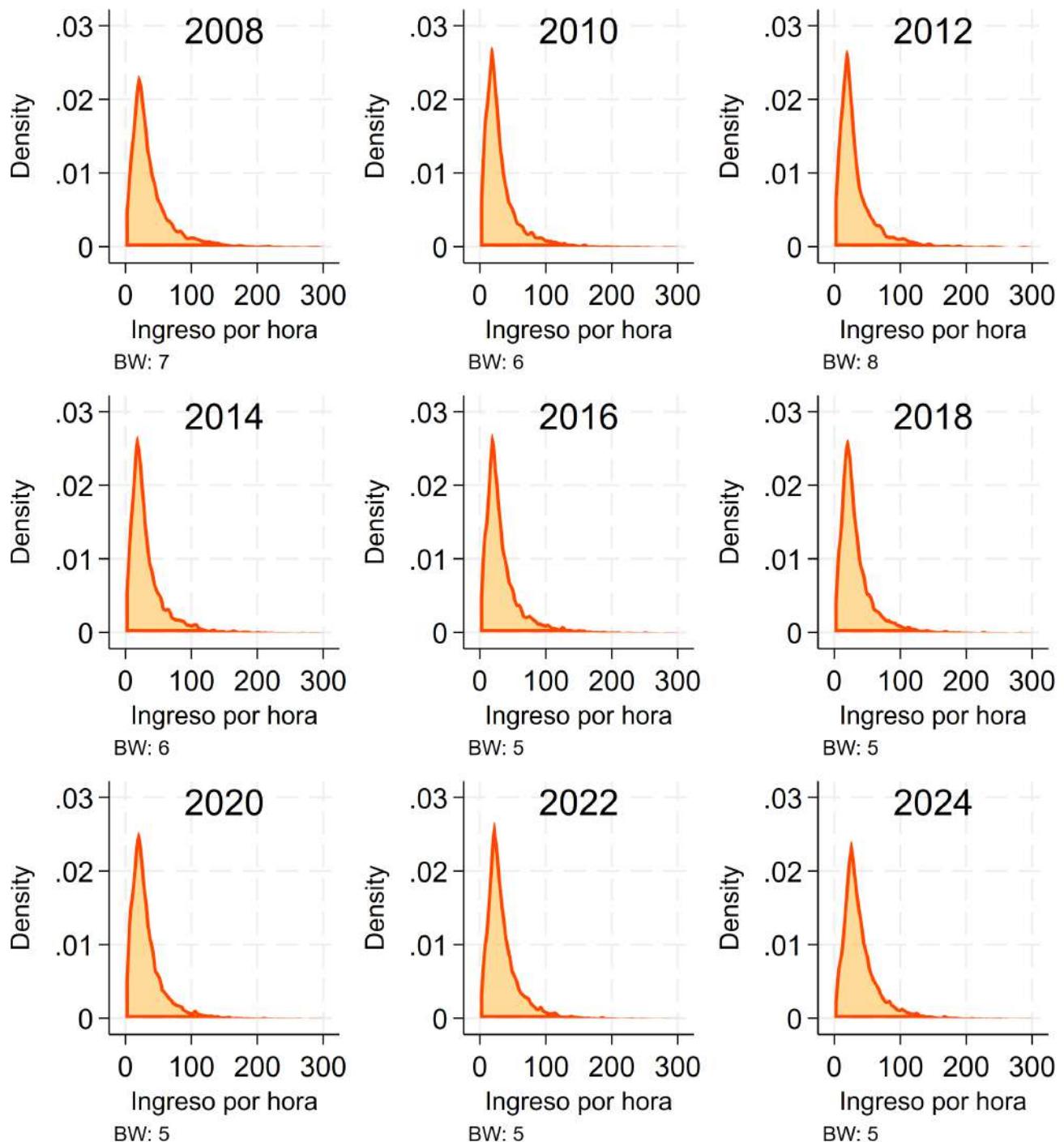
Al graficar la **distribución no paramétrica** del *salario por hora* con los datos de la **ENIGH 1992–2024**, puedo observar que la forma de las curvas se mantiene bastante similar a la de años previos. En todos los cortes el grueso de la masa salarial se concentra en valores *bajos*, lo que confirma la fuerte **asimetría** hacia la derecha. Me parece interesante cómo, a pesar del paso del tiempo, la *dispersión* y el *sesgo* se mantienen prácticamente inalterados, reflejando una **estructura laboral persistente**.

Figura 73: Estimaciones de densidad kernel del ingreso por hora (ENIGH 1992–2006)



Fuente: Elaboración propia con datos de la ENIGH. Proceso disponible en:  danifuentesga

Figura 74: Estimaciones de densidad kernel del ingreso por hora (ENIGH 2008–2024)



Fuente: Elaboración propia con datos de la ENIGH. Proceso disponible en:  danifuentesga

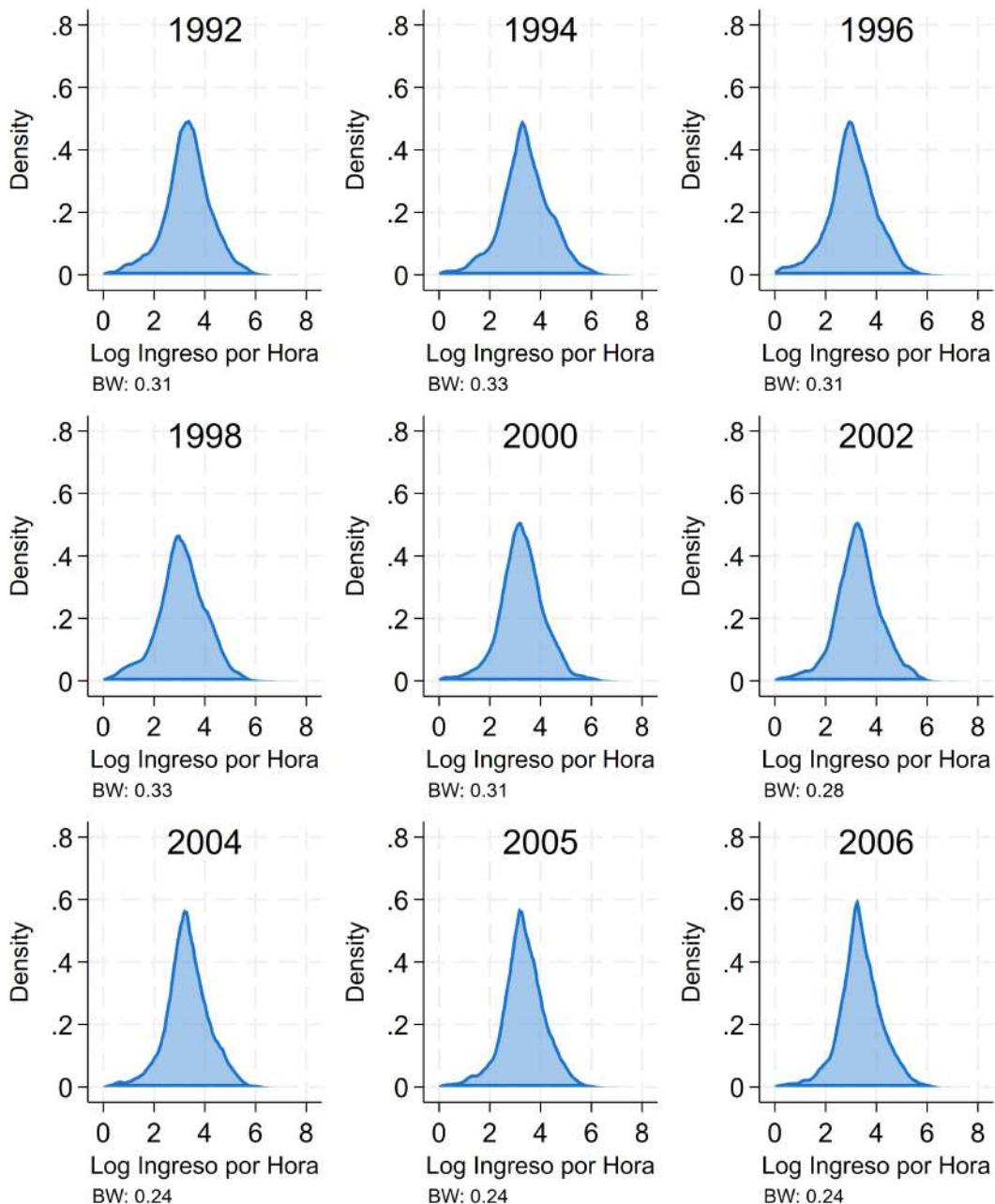
Ahora, genera una variable que sea el **logaritmo natural del salario por hora**. Grafica

una distribución no paramétrica para esta nueva variable.

¿Cómo cambian tus resultados?

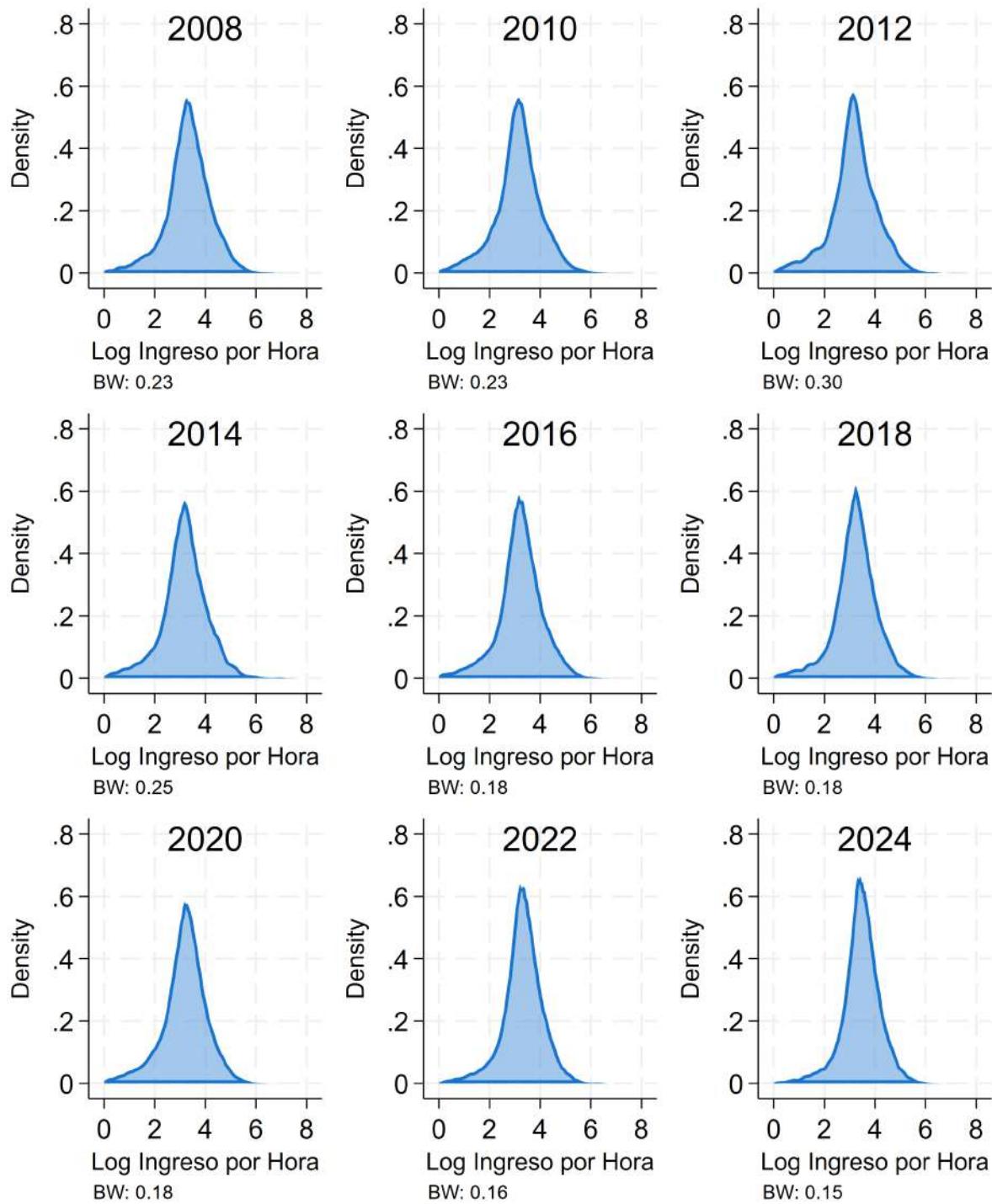
**RESPUESTA:**

Figura 75: Estimaciones de densidad kernel del logaritmo del ingreso por hora (ENIGH 1992–2006)



Fuente: Elaboración propia con datos de la ENIGH. Proceso disponible en:  danifuentesga

Figura 76: Estimaciones de densidad kernel del logaritmo del ingreso por hora (ENIGH 2008–2024)



Fuente: Elaboración propia con datos de la ENIGH. Proceso disponible en:  danifuentesga

Al transformar el ingreso por hora a su **logaritmo**, la distribución cambia de ser altamente *asimétrica a la derecha* (con una larga cola de pocos individuos con ingresos muy altos) a mostrar una forma mucho más **simétrica y concentrada**. En consecuencia, la curva se asemeja a una **distribución normal**, lo cual es útil porque facilita el análisis estadístico y la interpretación de diferencias relativas en los salarios.

### 5.3

Utilizando el **logaritmo natural del salario por hora**, crea dos **gráficas con distribuciones no paramétricas**:

- Una **aumentando** sustancialmente el **bandwidth**,
- Otra **disminuyéndolo**, utilizando un **único año**.

Haz lo mismo para **dos diferentes kernels**, manteniendo el **bandwidth constante**.

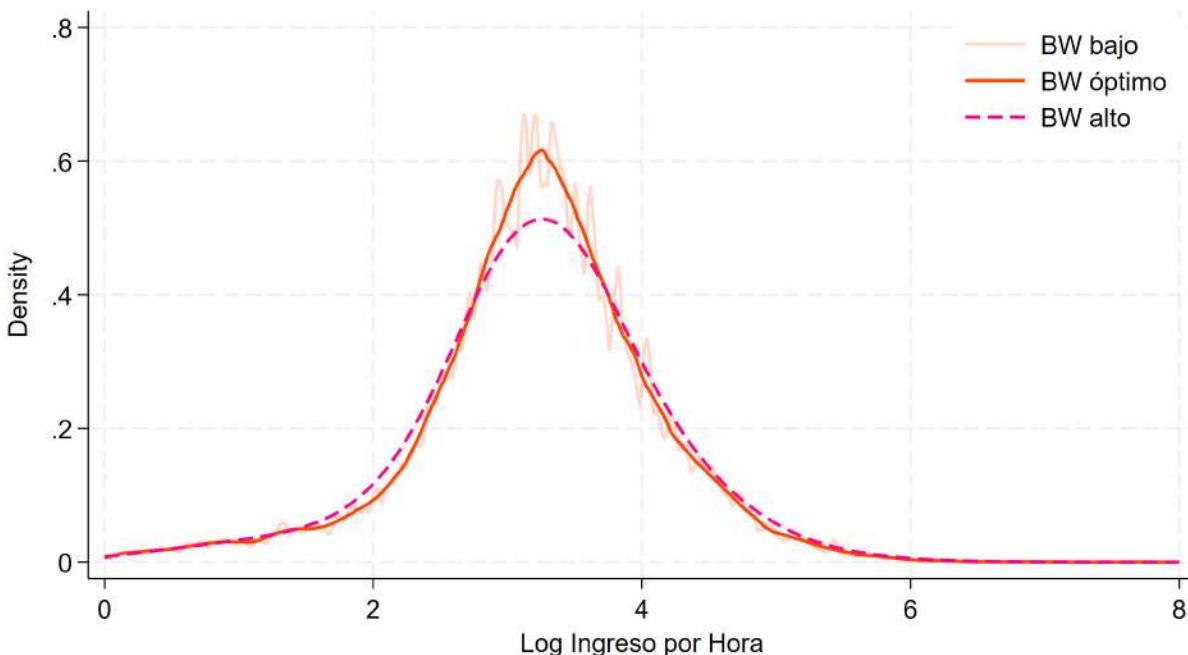
**¿Cómo afecta tu gráfica el bandwidth y el kernel? Explica.**

#### **RESPUESTA:**

El *bandwidth* controla el grado de suavización en la estimación de densidad kernel. Cuando el *bandwidth* es muy bajo, la curva se vuelve muy irregular y refleja demasiado el ruido de los datos. Con un *bandwidth* alto, la distribución se suaviza en exceso y puede ocultar rasgos importantes como asimetrías o múltiples modos. Un valor intermedio (óptimo) logra un balance, capturando la forma general de la distribución sin perder detalles relevantes.

El *bandwidth* óptimo en la estimación de densidad kernel se obtiene a partir de una regla de tipo Silverman, que depende de la desviación estándar y el tamaño muestral. La idea es escoger una ventana que minimice el error cuadrático medio, equilibrando el sesgo (que crece con un *bandwidth* grande) y la varianza (que crece con uno pequeño).

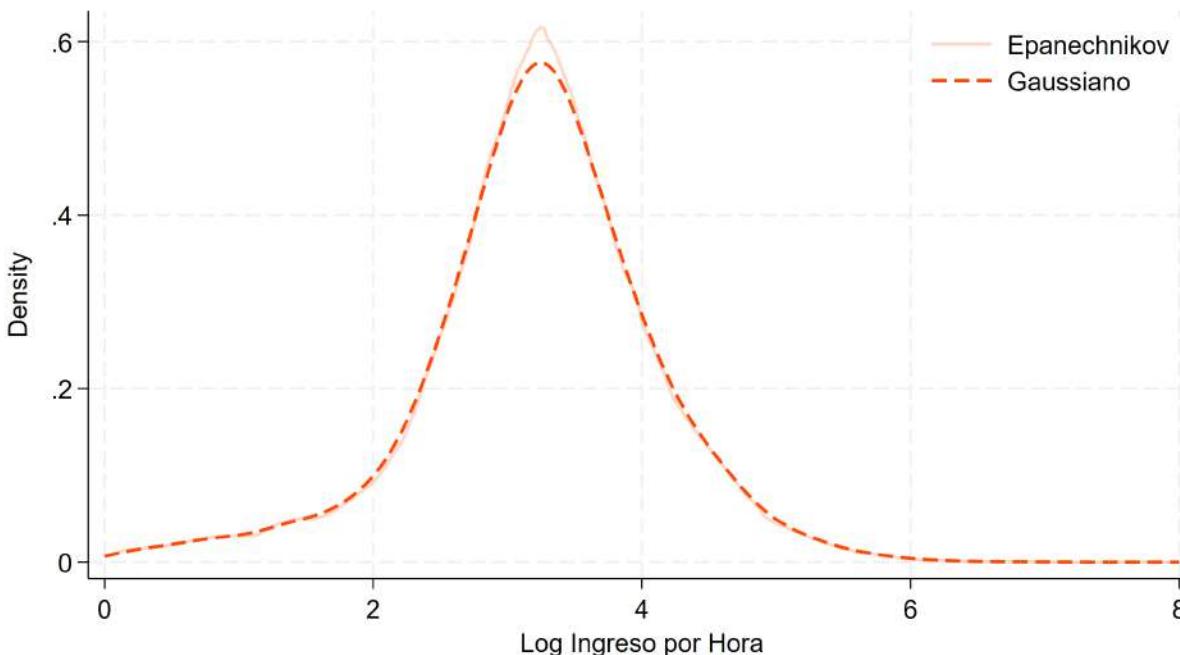
Figura 77: Comparación de estimaciones de densidad kernel del logaritmo del ingreso por hora bajo diferentes valores de bandwidth (ENIGH 2018)



Fuente: Elaboración propia con datos de la ENIGH. Proceso disponible en:  danifuentesga

La siguiente figura compara dos **funciones núcleo** (*Epanechnikov* y *Gaussiano*) en la **estimación de densidad del logaritmo del salario por hora**. Se observa que ambas producen curvas **muy similares** en forma y tendencia, con un **pico principal** alrededor del valor modal y colas que decrecen suavemente. La diferencia clave es que el *Gaussiano* genera una curva más **redondeada**, mientras que el *Epanechnikov* resulta ligeramente más **afilado**, aunque el impacto visual es **mínimo**.

Figura 78: Comparación de estimaciones de densidad kernel con diferentes funciones kernel (ENIGH 2018)



Fuente: Elaboración propia con datos de la ENIGH. Proceso disponible en:  danifuentesga

## 5.4

Grafica regresiones no paramétricas utilizando la ENIGH.

a)

Explica las diferencias entre `lowess` y `lpoly`

### RESPUESTA:

La diferencia principal es que `lowess` (*locally weighted scatterplot smoothing*) es un método **no paramétrico** que ajusta **rectas locales** a los datos, asignando mayor peso a los puntos cercanos en cada vecindad. Por otro lado, `lpoly` (*local polynomial regression*) generaliza esta idea y permite ajustar **polinomios locales** (no sólo rectas), lo que le da más flexibilidad para capturar *curvaturas* o patrones complejos en la nube de datos. En resumen:

**lowess** = rectas locales, mientras que **lpoly** = polinomios locales.

b)

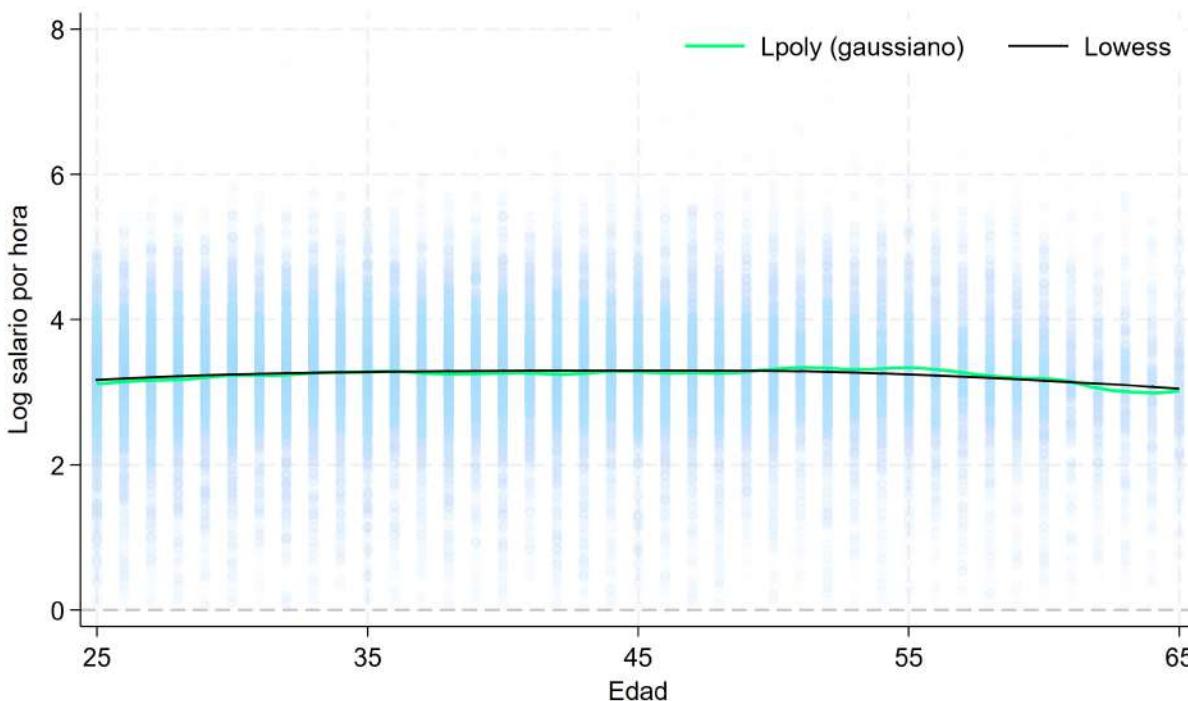
Escoge el año de **2018**. Grafica no paramétricamente utilizando lowess y lpoly el **logaritmo del salario por hora**, separadamente para **hombres y mujeres**.

En el eje y va el salario. Realiza dos figuras:

- una donde el eje x sea la **edad**,
- otra donde el eje x sean los **años de escolaridad**.

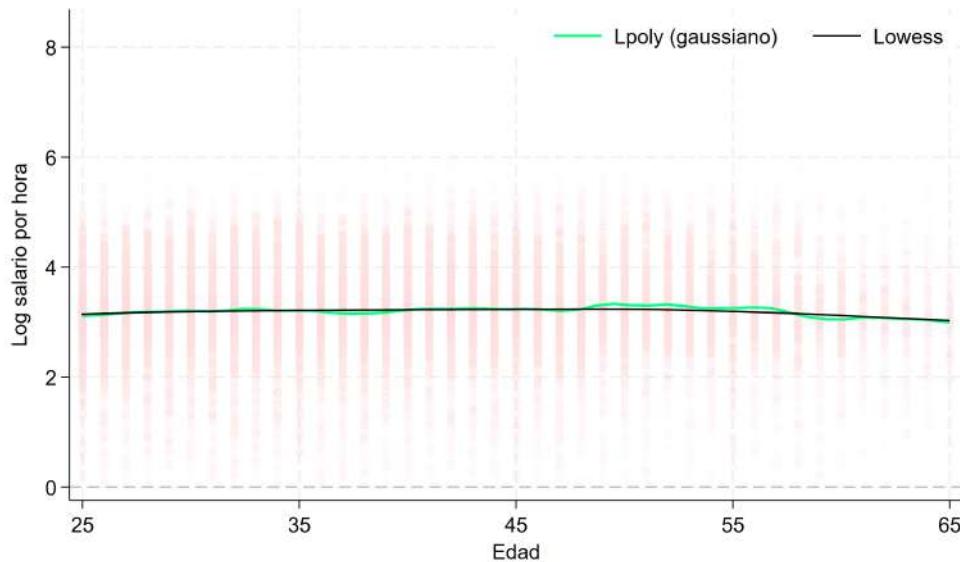
### **RESPUESTA:**

Figura 79: Relación no paramétrica entre edad y logaritmo del salario por hora en hombres (ENIGH 2018)



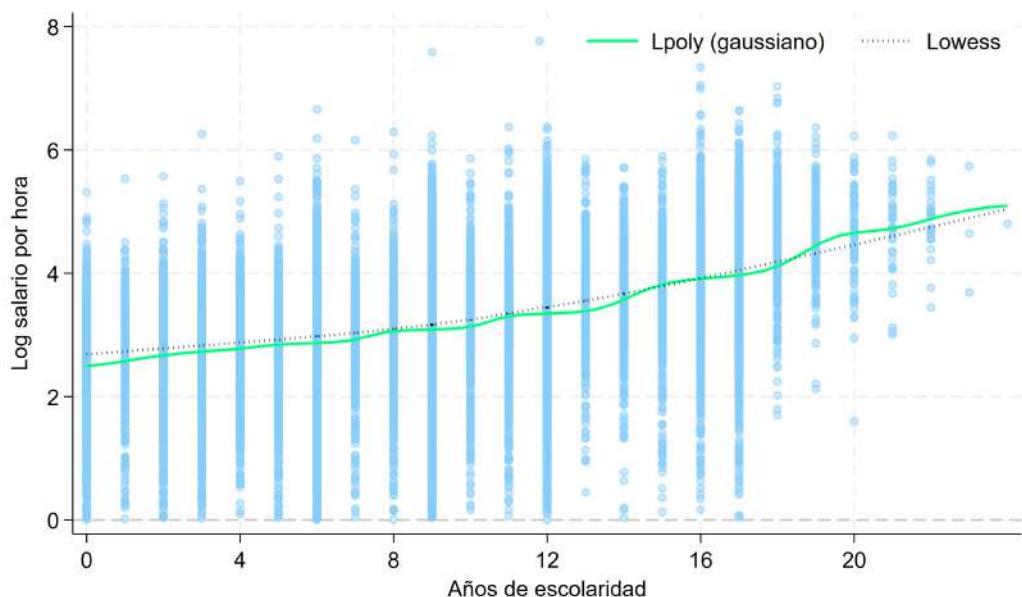
Fuente: Elaboración propia con datos de la ENIGH. Proceso disponible en:  danifuentesga

Figura 80: Relación no paramétrica entre edad y logaritmo del salario por hora en mujeres (ENIGH 2018)



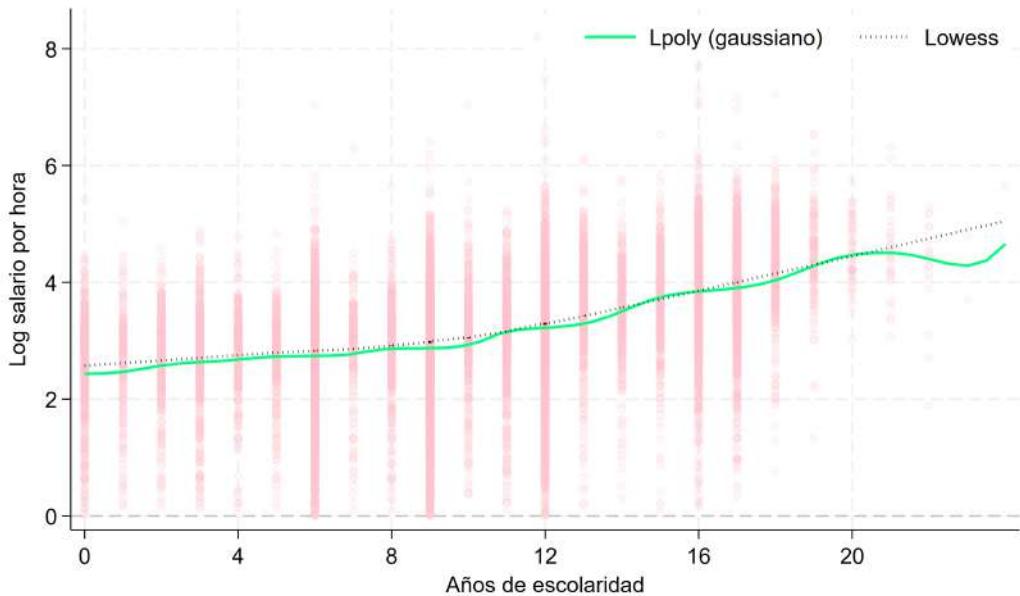
Fuente: Elaboración propia con datos de la ENIGH. Proceso disponible en:  danifuentesga

Figura 81: Relación no paramétrica entre escolaridad y logaritmo del salario por hora en hombres (ENIGH 2018)



Fuente: Elaboración propia con datos de la ENIGH. Proceso disponible en:  danifuentesga

Figura 82: Relación no paramétrica entre escolaridad y logaritmo del salario por hora en mujeres (ENIGH 2018)



Fuente: Elaboración propia con datos de la ENIGH. Proceso disponible en:  danifuentesga

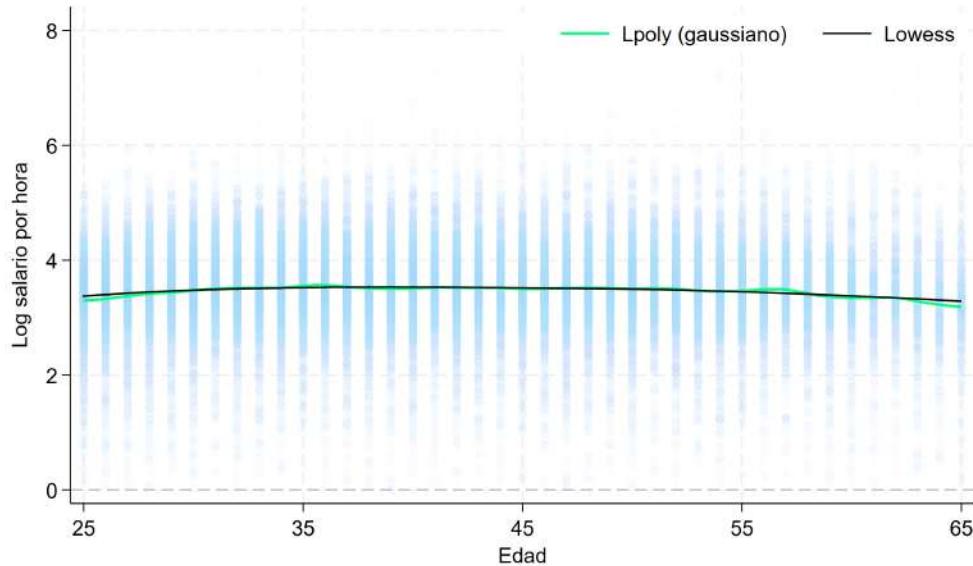
Las estimaciones no paramétricas con `lowess` y `lpoly` permiten observar de manera flexible la relación entre el **logaritmo del salario por hora** y las variables de interés. Cuando el eje  $x$  es la **edad**, tanto en hombres como en mujeres, la curva es prácticamente plana, lo cual sugiere que, controlando de manera no paramétrica, la edad no genera grandes variaciones en el salario horario. En cambio, cuando el eje  $x$  son los **años de escolaridad**, la pendiente es positiva: a mayor escolaridad, mayor es el log salario por hora, lo que confirma la relevancia del capital humano. El supuesto clave es que los métodos `lowess` y `lpoly` no imponen una forma funcional rígida, sino que suavizan los datos localmente; por ello, la interpretación depende de que asumimos que los salarios siguen una relación suave con la edad o la educación, sin saltos bruscos.

c)

Escoge el año de 2024. Repite el ejercicio b.

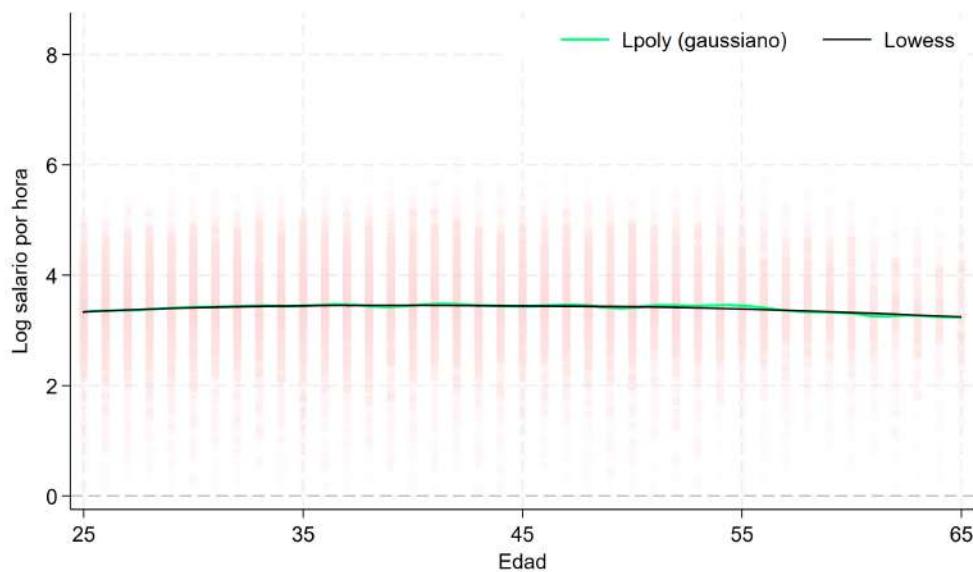
## RESPUESTA:

Figura 83: Relación no paramétrica entre edad y logaritmo del salario por hora en hombres (ENIGH 2024)



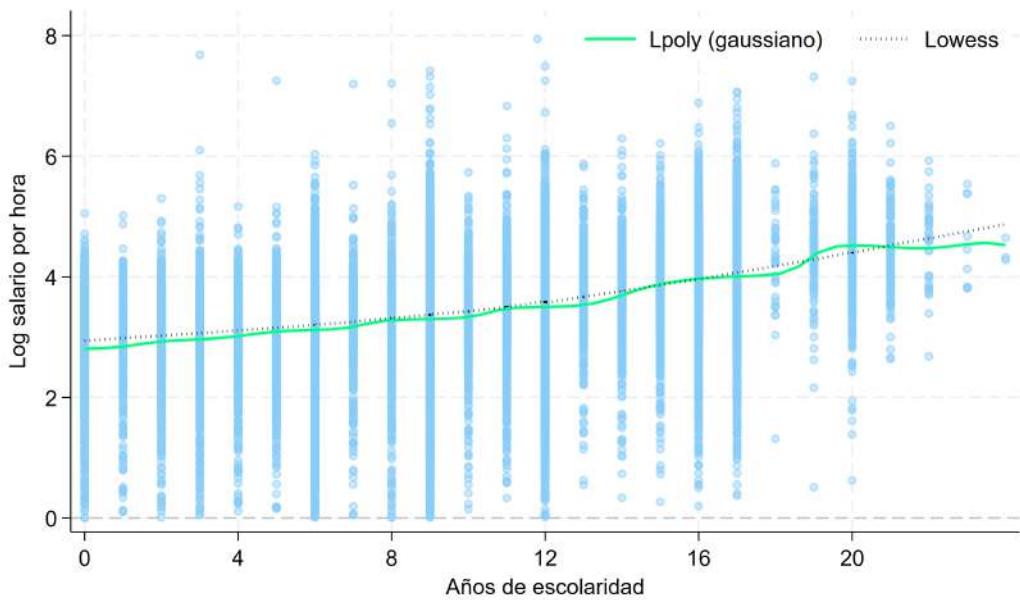
Fuente: Elaboración propia con datos de la ENIGH. Proceso disponible en:  danifuentesga

Figura 84: Relación no paramétrica entre edad y logaritmo del salario por hora en mujeres (ENIGH 2024)



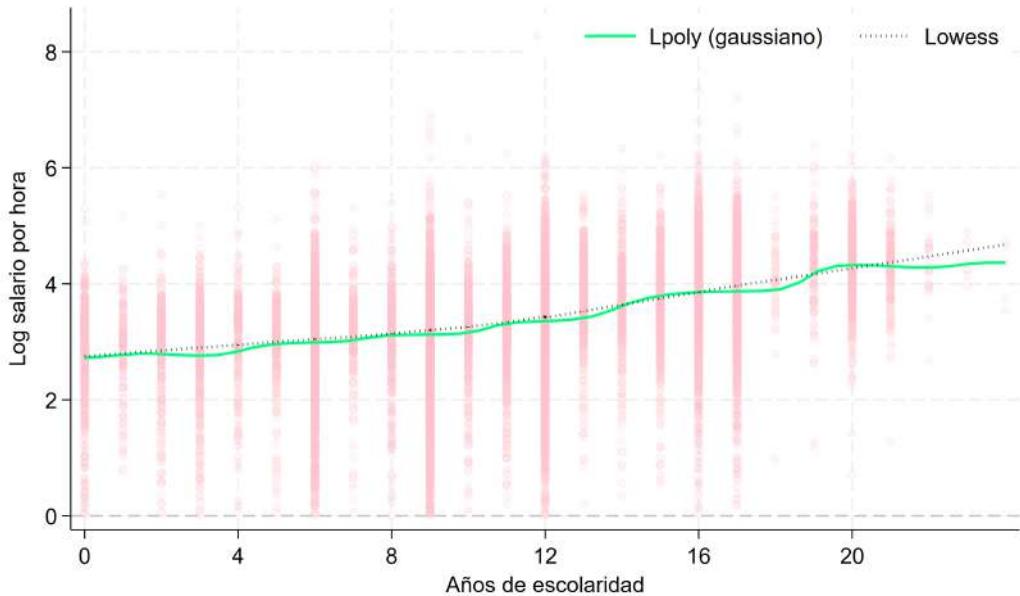
Fuente: Elaboración propia con datos de la ENIGH. Proceso disponible en:  danifuentesga

Figura 85: Relación no paramétrica entre escolaridad y logaritmo del salario por hora en hombres (ENIGH 2024)



Fuente: Elaboración propia con datos de la ENIGH. Proceso disponible en:  danifuentesga

Figura 86: Relación no paramétrica entre escolaridad y logaritmo del salario por hora en mujeres (ENIGH 2024)



Fuente: Elaboración propia con datos de la ENIGH. Proceso disponible en:  danifuentesga

d)

**Explica diferencias y similitudes. Explica posibles razones.**

*¿Crees que esto nos ayude a explicar qué pasó con el ingreso laboral en esos dos años en términos de comparabilidad?*

**RESPUESTA:**

Las estimaciones no paramétricas muestran patrones diferentes según la variable en el eje  $x$ . Cuando se usa la **edad**, tanto en 2018 como en 2024 las curvas (`lowess`, `lpoly`) son casi planas, lo cual indica que la edad por sí sola no explica de manera significativa el **logaritmo del salario por hora**. Esta similitud refuerza la idea de que la edad no es un buen predictor del salario horario en ninguno de los dos años, aunque en 2024 se percibe una leve caída en edades avanzadas.

Por otro lado, cuando el eje  $x$  son los **años de escolaridad**, en ambos años se observa una relación claramente positiva: a mayor escolaridad, mayor salario horario en logaritmos. Sin embargo, la **diferencia** es que en 2018 la pendiente parece más pronunciada en niveles altos de educación, mientras que en 2024 la curva es más suave, sugiriendo una *possible disminución en los retornos de la educación*.

En cuanto a la **comparabilidad**, los resultados deben tomarse con cautela. Si bien las formas generales son consistentes (curva plana para edad y creciente para educación), las diferencias en la magnitud de las pendientes podrían deberse tanto a cambios reales en el mercado laboral como a variaciones en la captación de los ingresos por parte de la ENIGH. Por tanto, aunque las gráficas aportan *pistas valiosas*, no garantizan una plena comparabilidad entre 2018 y 2024 sin considerar otros factores que podrían ser relevantes

e)

Un problema de las gráficas `lowess` y `lpoly` es que no controlan por ninguna **variable explicativa**.

Existe **análisis semiparamétrico** que permite controlar por variables explicativas. (*Puedes leer el libro de Cameron y Trivedi para una mayor guía.*)

- i) Explica qué es análisis semiparamétrico.

**RESPUESTA:**

Un **análisis semiparamétrico** combina dos aspectos: la parte **paramétrica**, donde se impone una estructura funcional conocida (por ejemplo, una relación lineal en algunos regresores), y la parte **no paramétrica**, donde se deja que los datos “hablen” libremente sin imponer una forma rígida. Formalmente, un modelo semiparamétrico puede escribirse como:

$$E[y|x, z] = x'\beta + g(z),$$

donde  $x$  son regresores tratados de forma *paramétrica* (con coeficientes  $\beta$  que interpretamos directamente) y  $z$  entra de manera *no paramétrica* a través de una función  $g(\cdot)$  desconocida. De esta forma, podemos controlar por ciertas variables al estilo regresión clásica, mientras permitimos flexibilidad en otras.

Aterrizando a nuestro caso, si modelamos el **logaritmo del salario por hora**, podríamos incluir de forma paramétrica variables de control como **sexo, región o sector laboral** (que tienen efectos más claros y lineales), y dejar de manera no paramétrica variables como la **edad** o los **años de escolaridad**, cuya relación con el salario hemos visto que es curva y difícil de capturar con una recta simple. Así, el análisis semiparamétrico supera la limitación de `lowess` y `lpoly`, pues permite analizar la relación *suave* entre educación o edad y el salario, pero al mismo tiempo controlando por otros factores importantes. En este sentido, aporta un marco más realista y útil para interpretar los retornos a la educación y la evolución de los ingresos laborales.

- ii) Asimismo, si tomas una **regresión de salario vs variables explicativas**, y de **edad vs variables explicativas**, tomas los *residuales* y graficas con `lowess` y `lpoly`, podrías

solucionar ese problema.

Haz ese cálculo para los años **2018** y **2022**, y explica la *intuición del mecanismo*.

**Discute tus resultados.** Es decir, grafica la **relación entre salario y edad**, *controlando por diferentes variables explicativas*.

**RESPUESTA:**

La idea es **aislar la relación “pura” entre edad y salario**, eliminando el efecto de variables de control que también influyen: **sexo, rural y años de escolaridad**.

- **Primer paso (Reg1):** corremos la regresión

$$\ln(wage) = \alpha + \gamma_1 \text{sexo} + \gamma_2 \text{rural} + \gamma_3 \text{escolaridad} + u$$

y guardamos el residual  $\hat{u}$ . Ese residual se interpreta como el “*salario neto de controles*”, es decir, el componente del logaritmo del salario que no puede ser explicado por sexo, ruralidad ni escolaridad.

- **Segundo paso (Reg2):** corremos la regresión

$$edad = \delta + \theta_1 \text{sexo} + \theta_2 \text{rural} + \theta_3 \text{escolaridad} + v$$

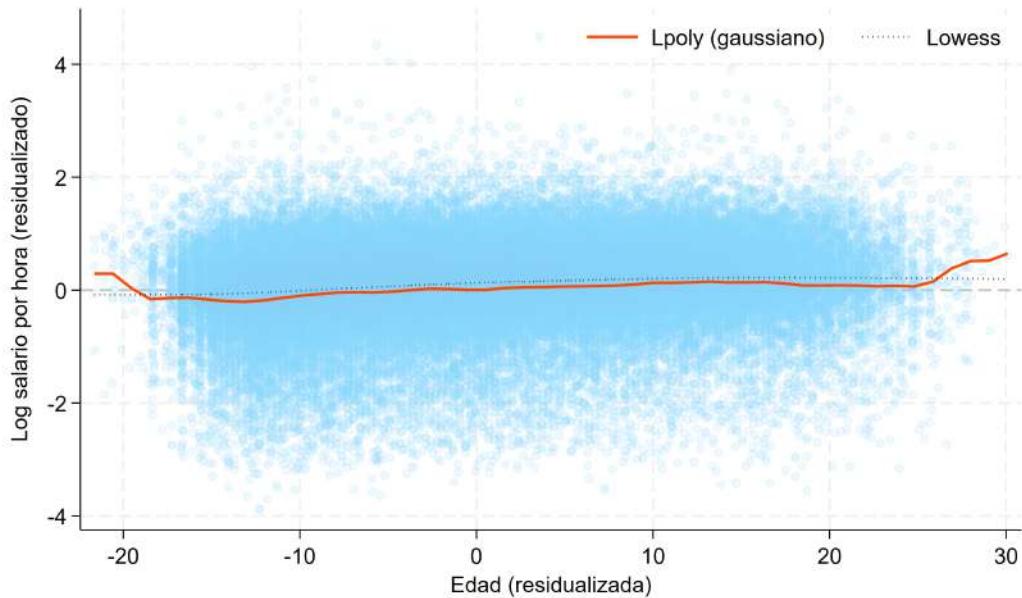
y guardamos el residual  $\hat{v}$ . Este residual se interpreta como la “*edad neta de controles*”, es decir, la parte de la edad que no está correlacionada con esas tres características.

- **Tercer paso:** graficamos  $\hat{u}$  contra  $\hat{v}$  utilizando métodos no paramétricos (`lowess` y `lpoly`).

Al graficar los residuales  $\hat{u}$  y  $\hat{v}$ , obtenemos la *relación directa entre edad y salario*, pero ya **controlando por sexo, escolaridad y ruralidad**. La lógica es que al quitar de **salario** y de **edad** la parte explicada por **sexo, rural y escolaridad**, lo que queda en los residuales  $\hat{u}$  y  $\hat{v}$  es la variación que no depende de esos controles. Si graficamos  $\hat{u}$  contra  $\hat{v}$ , vemos

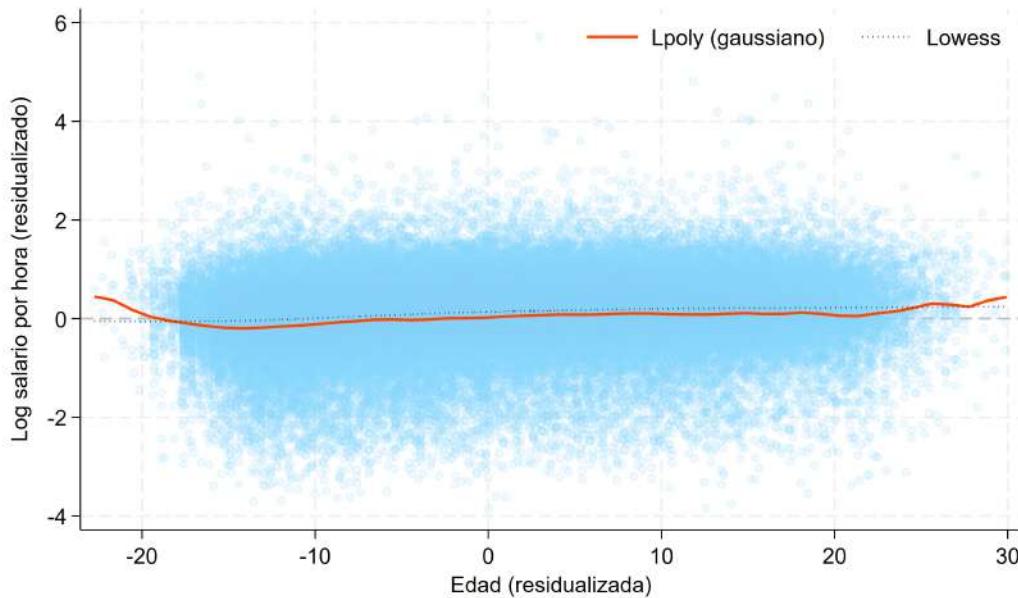
únicamente cómo se relacionan salario y edad en su componente no explicado por otras variables. En otras palabras: **los controles actúan como un filtro**, y la curva residual captura la asociación neta entre edad y salario.

Figura 87: Relación neta entre edad y logaritmo del salario por hora, controlando por escolaridad, sexo y ruralidad (ENIGH,2018)



Fuente: Elaboración propia con datos de la ENIGH. Proceso disponible en:  danifuentesga

Figura 88: Relación neta entre edad y logaritmo del salario por hora, controlando por escolaridad, sexo y ruralidad (ENIGH,2022)



Fuente: Elaboración propia con datos de la ENIGH. Proceso disponible en:  danifuentesga

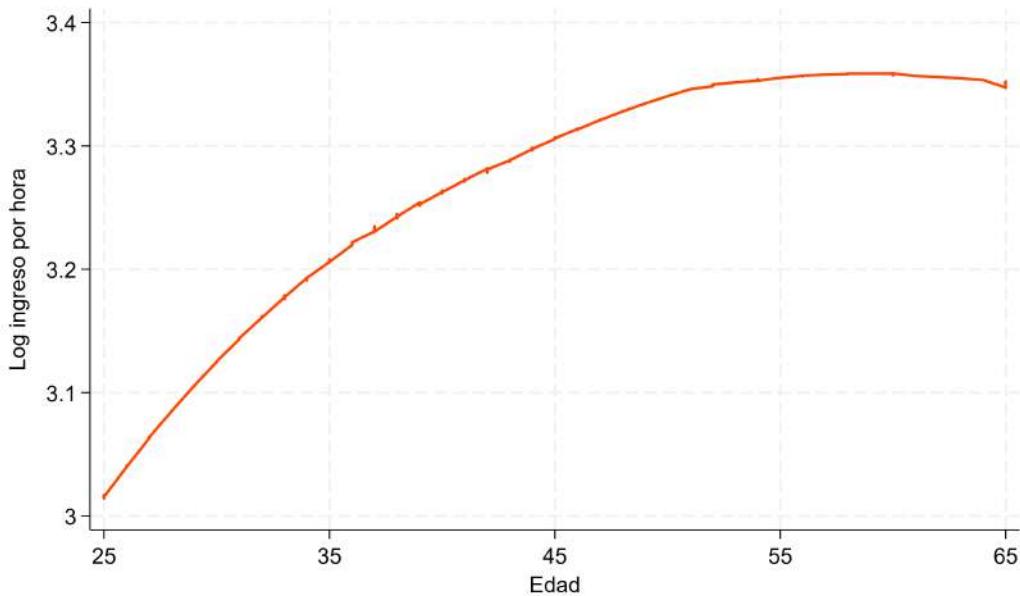
- iii) Baja el ado file `plreg`, checa el `help`, y compara tus **gráficas y resultados anteriores** con el **resultado de ese comando** también.

**Discute.**

### **RESPUESTA:**

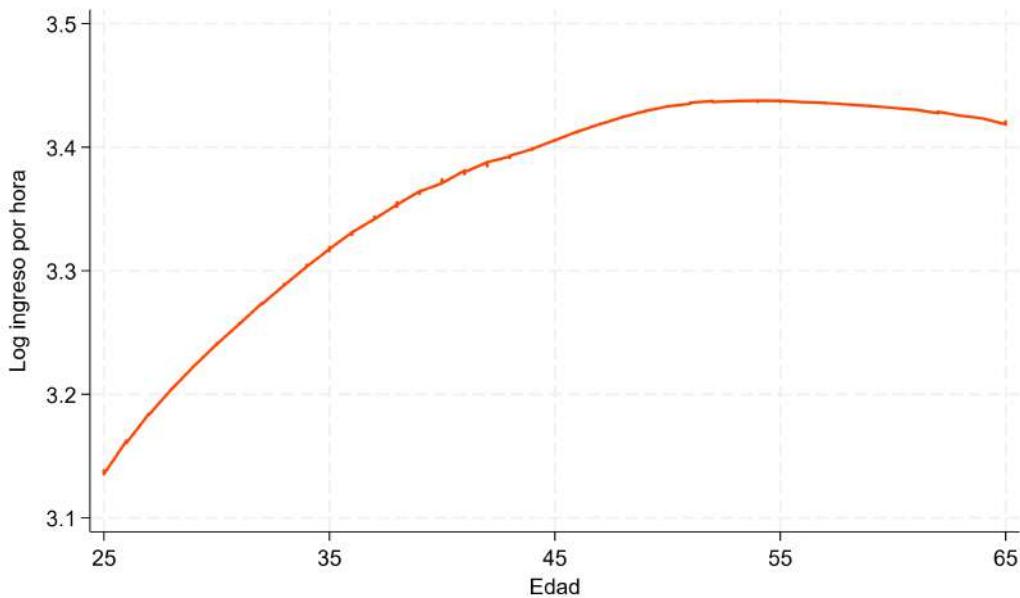
En pocas palabras, las dos figuras responden a preguntas distintas. La de *residuales* (con `lowess/lpoly`) muestra la relación **neta** entre edad y *log salario* tras “limpiar” el efecto de **sexo, rural y años de escolaridad**; por eso la curva luce casi plana: una vez controlados esos factores, la edad agrega poco. En cambio, `plreg` estima un modelo **semiparamétrico** donde la edad entra de forma no paramétrica y los controles lo hacen linealmente, revelando crecimiento del salario con la edad. Así, la primera gráfica es ideal para interpretar el **efecto puramente parcial** de la edad, mientras que `plreg` resume la **trayectoria suavizada** del salario por edad, manteniendo los controles pero sin “des-residualizar” la variable dependiente.

Figura 89: Relación semiparamétrica entre edad e ingreso laboral horario en logaritmos (ENIGH, 2018)



Fuente: Elaboración propia con datos de la ENIGH. Proceso disponible en:  danifuentesga

Figura 90: Relación semiparamétrica entre edad e ingreso laboral horario en logaritmos (ENIGH, 2022)

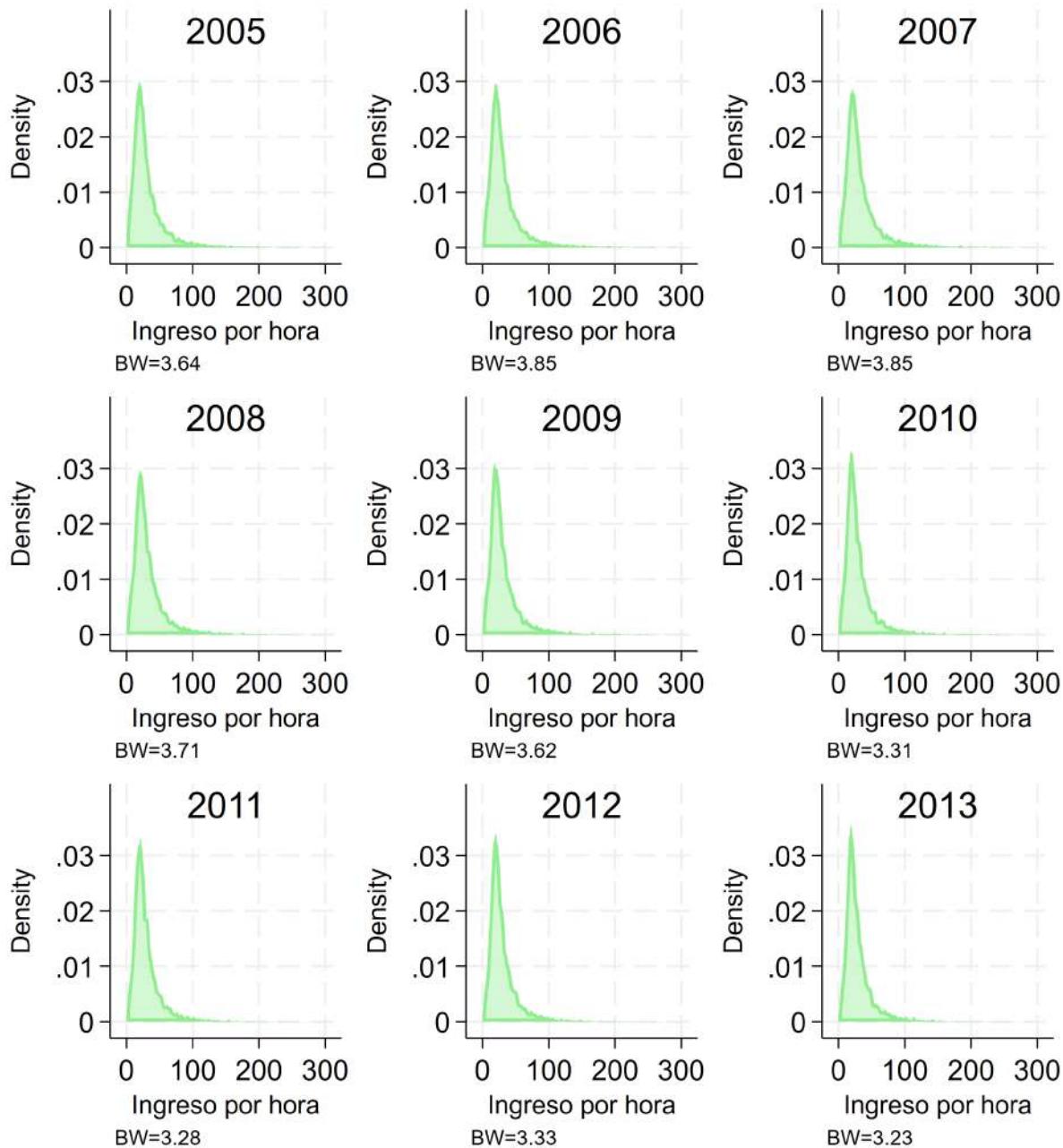


Fuente: Elaboración propia con datos de la ENIGH. Proceso disponible en:  danifuentesga

## 5.5

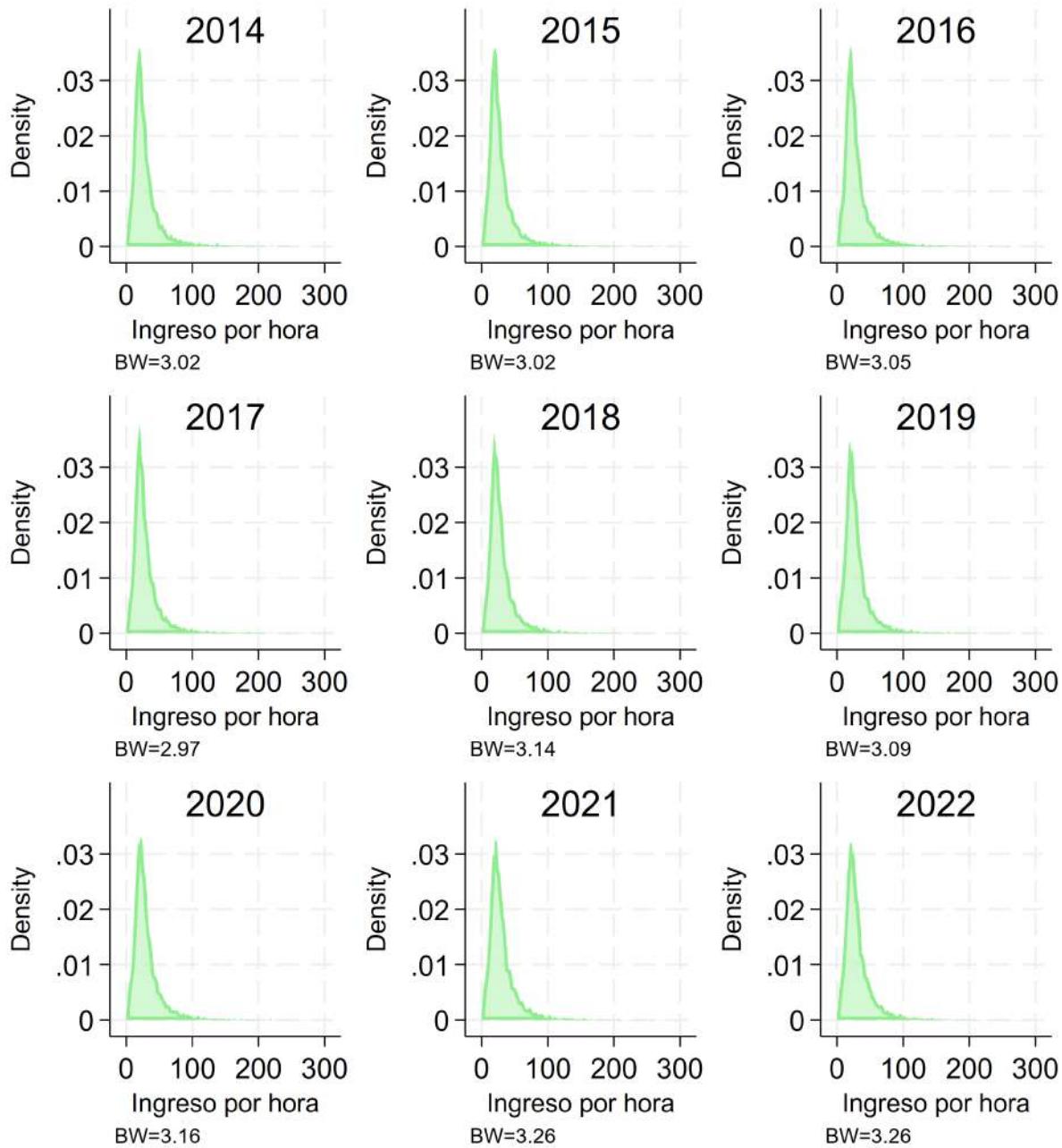
Repite el ejercicio con ENOE para los mismos años (*usar trimestre II*), y discute la **comparabilidad entre ENIGH y ENOE**.

Figura 91: Estimaciones de densidad kernel del ingreso por hora (ENOE, 2005–2013)



Fuente: Elaboración propia con datos de la ENOE. Proceso disponible en:  danifuentesga

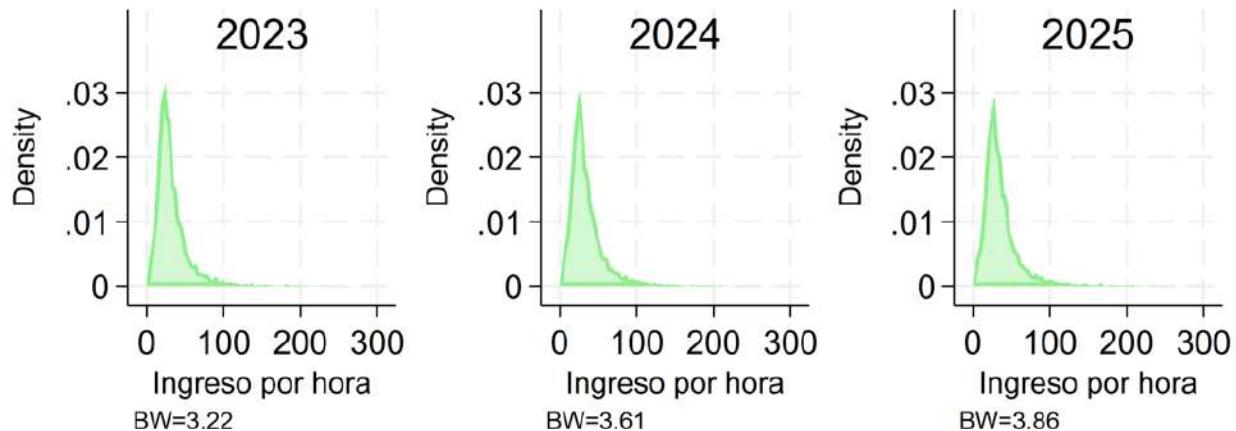
Figura 92: Estimaciones de densidad kernel del ingreso por hora (ENOE, 2014–2022)



*Nota:* Todas las estimaciones utilizan información del trimestre II de cada año, excepto el año 2020, donde se empleó el trimestre I.

Fuente: Elaboración propia con datos de la ENOE. Proceso disponible en:  danifuentesga

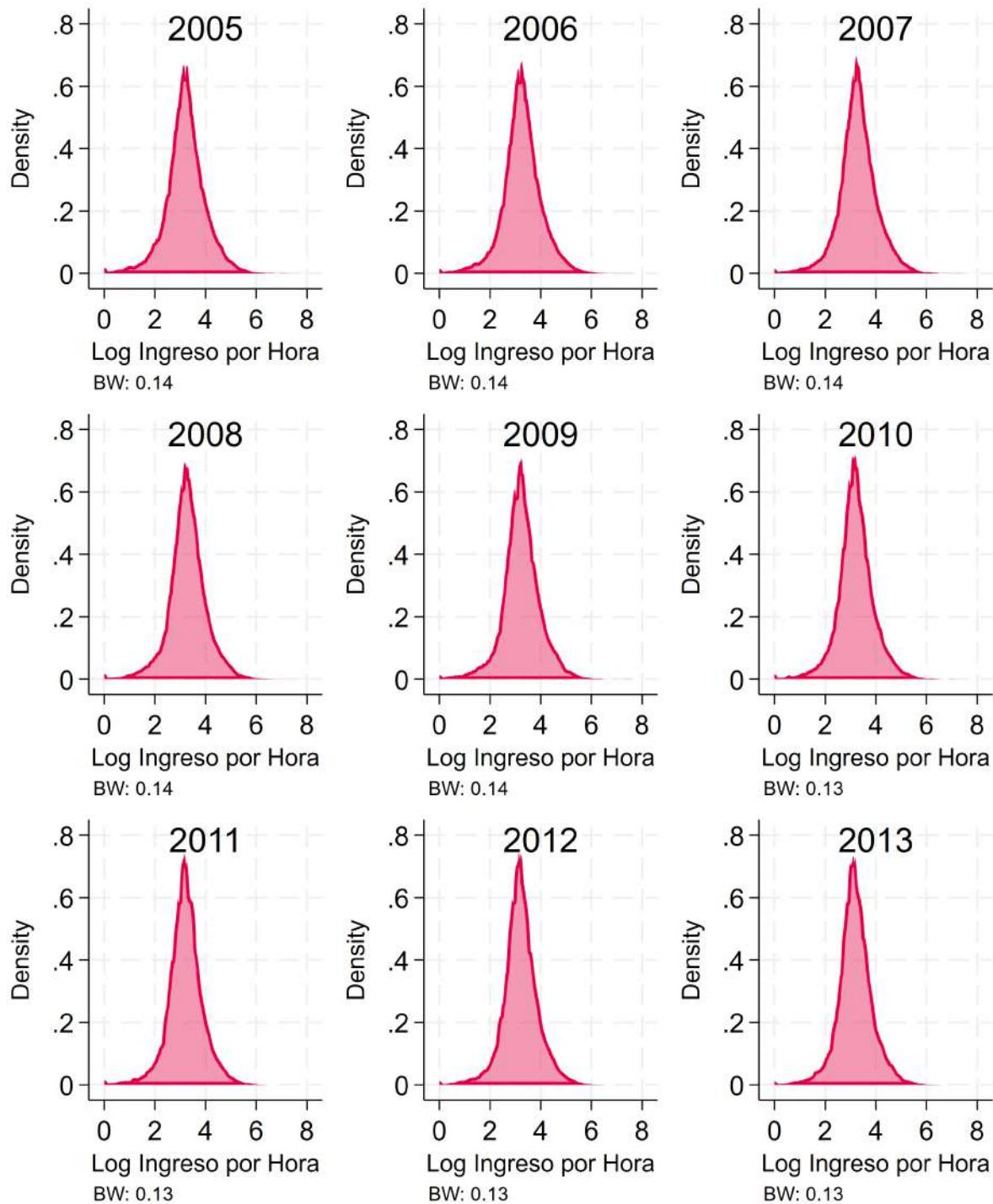
Figura 93: Estimaciones de densidad kernel del ingreso por hora (ENOE, 2023–2025)



*Nota:* Todas las estimaciones utilizan información del trimestre II de cada año, excepto el año 2020, donde se empleó el trimestre I.

Fuente: Elaboración propia con datos de la ENOE. Proceso disponible en:  danifuentesga

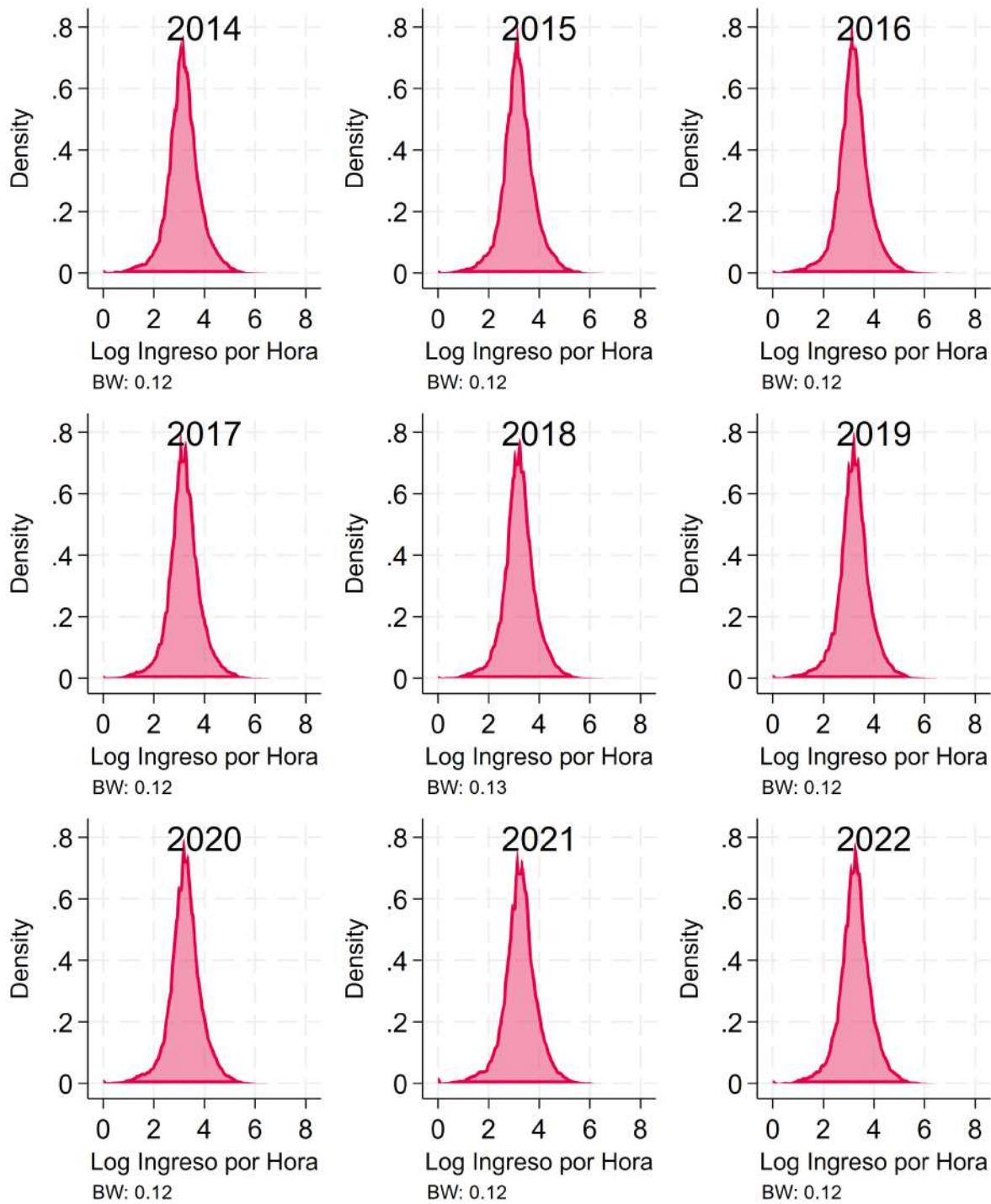
Figura 94: Estimaciones de densidad kernel del logaritmo del ingreso por hora (ENOE, 2005–2013)



*Nota:* Todas las estimaciones utilizan información del trimestre II de cada año, excepto el año 2020, donde se empleó el trimestre I.

Fuente: Elaboración propia con datos de la ENOE. Proceso disponible en:  danifuentesga

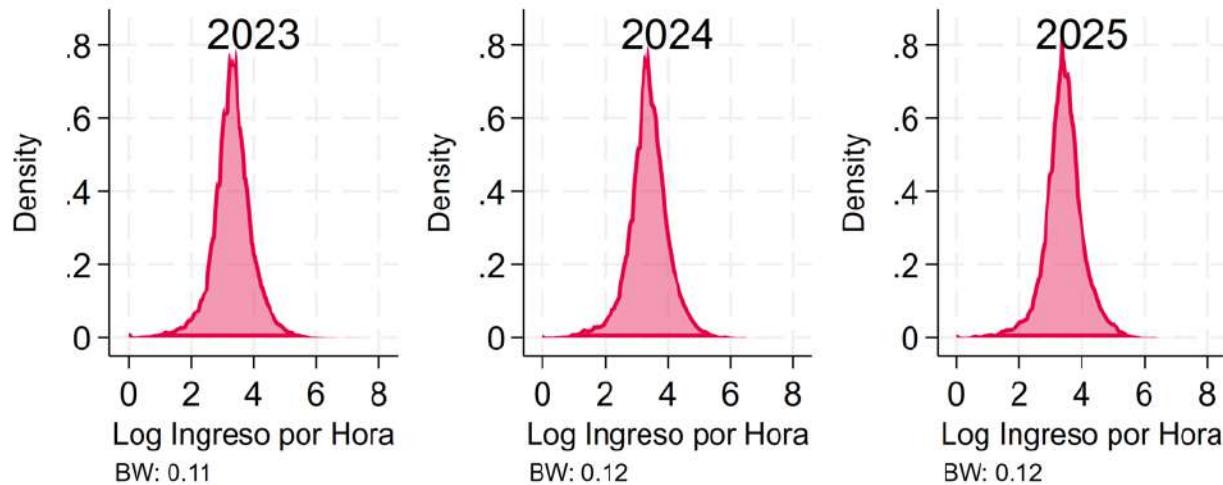
Figura 95: Estimaciones de densidad kernel del logaritmo del ingreso por hora (ENOE, 2014–2022)



*Nota:* Todas las estimaciones utilizan información del trimestre II de cada año, excepto el año 2020, donde se empleó el trimestre I.

Fuente: Elaboración propia con datos de la ENOE. Proceso disponible en:  danifuentesga

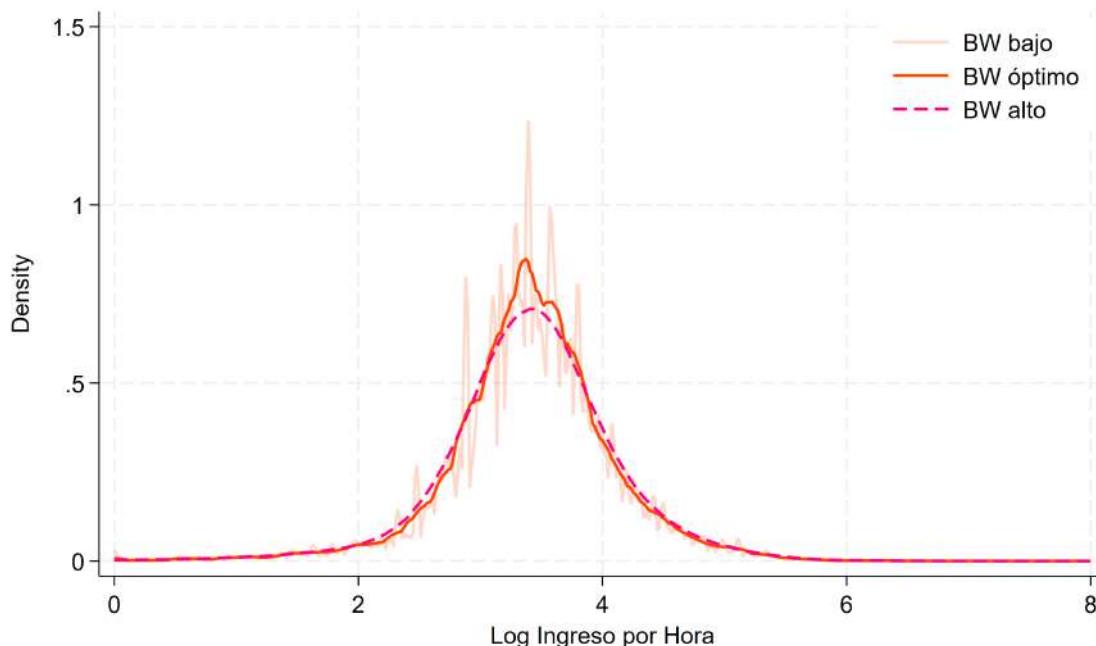
Figura 96: Estimaciones de densidad kernel del logaritmo del ingreso por hora (ENOE, 2023–2025)



*Nota:* Todas las estimaciones utilizan información del trimestre II de cada año, excepto el año 2020, donde se empleó el trimestre I.

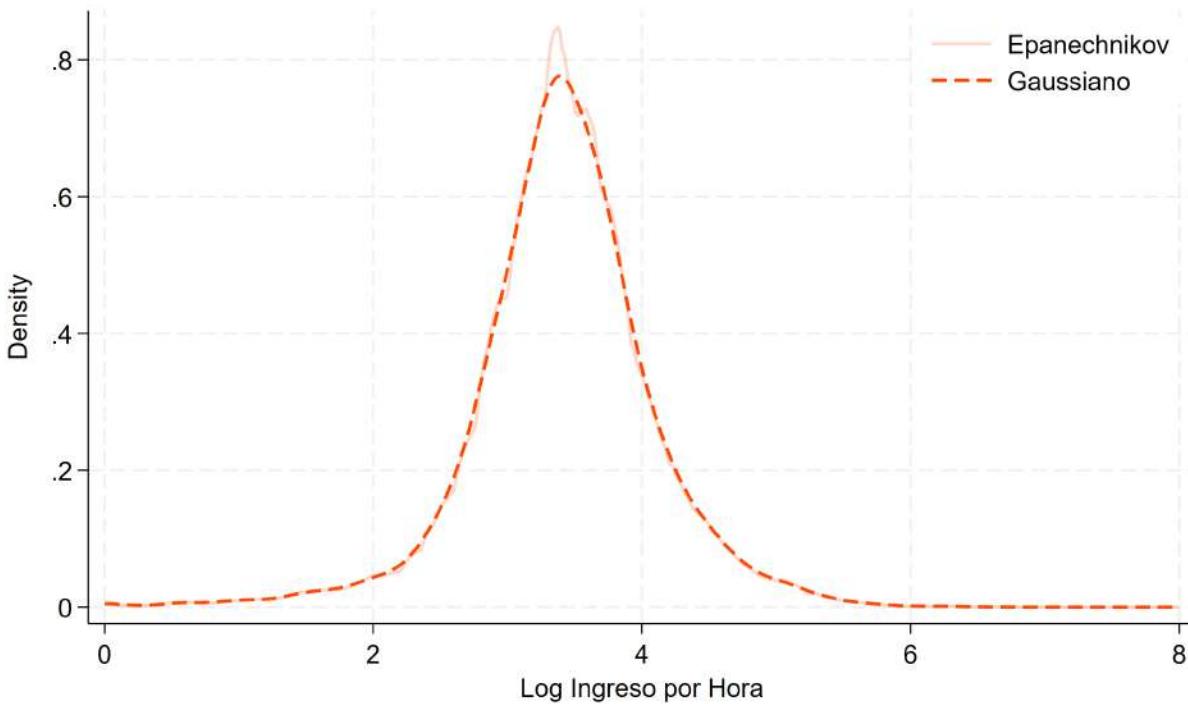
Fuente: Elaboración propia con datos de la ENOE. Proceso disponible en:  danifuentesga

Figura 97: Comparación de estimaciones de densidad kernel del logaritmo del ingreso por hora bajo diferentes valores de bandwidth (ENOE 2025)



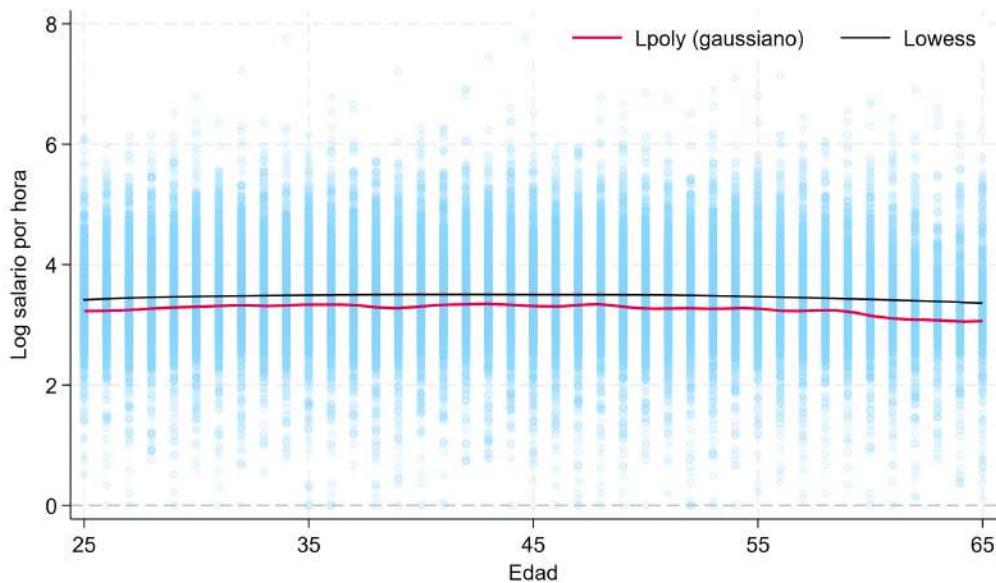
Fuente: Elaboración propia con datos de la ENOE. Proceso disponible en:  danifuentesga

Figura 98: Comparación de estimaciones de densidad kernel con diferentes funciones kernel (ENOE, 2025)



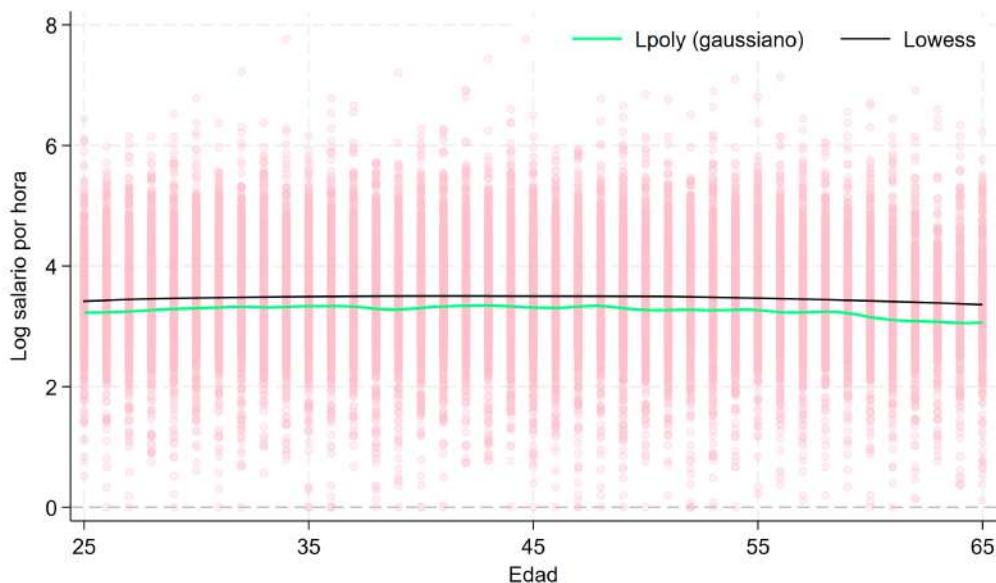
Fuente: Elaboración propia con datos de la ENOE. Proceso disponible en:  danifuentesga

Figura 99: Relación no paramétrica entre edad y logaritmo del salario por hora en hombres (ENOE 2018)



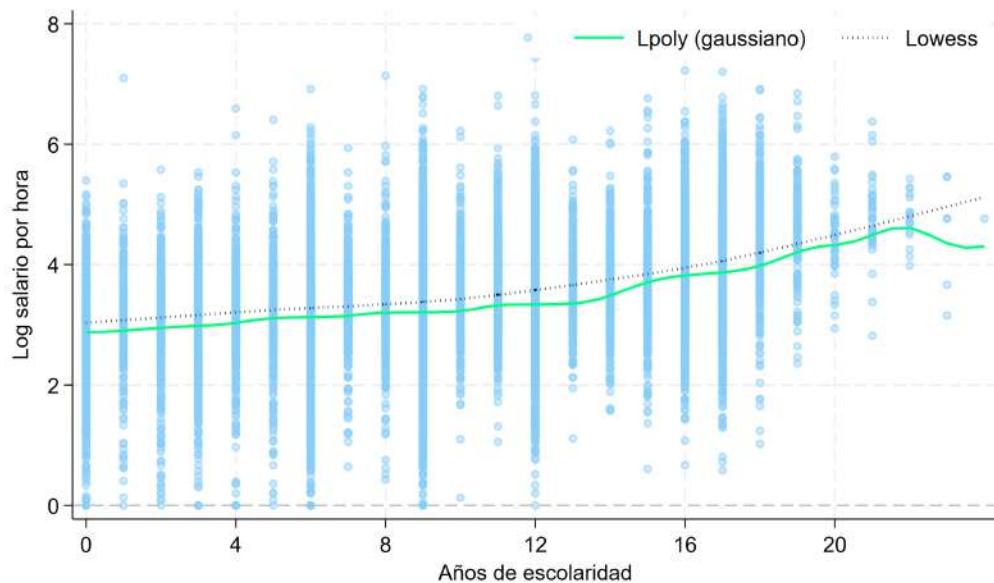
Fuente: Elaboración propia con datos de la ENOE. Proceso disponible en:  danifuentesga

Figura 100: Relación no paramétrica entre edad y logaritmo del salario por hora en mujeres (ENOE 2018)



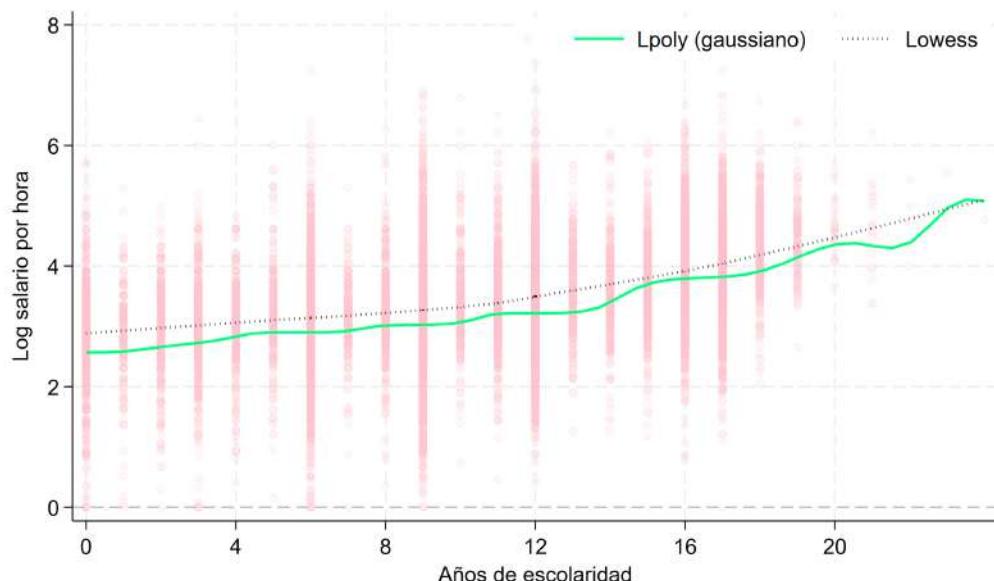
Fuente: Elaboración propia con datos de la ENOE. Proceso disponible en:  danifuentesga

Figura 101: Relación no paramétrica entre esolaridad y logaritmo del salario por hora en hombres (ENOE 2018)



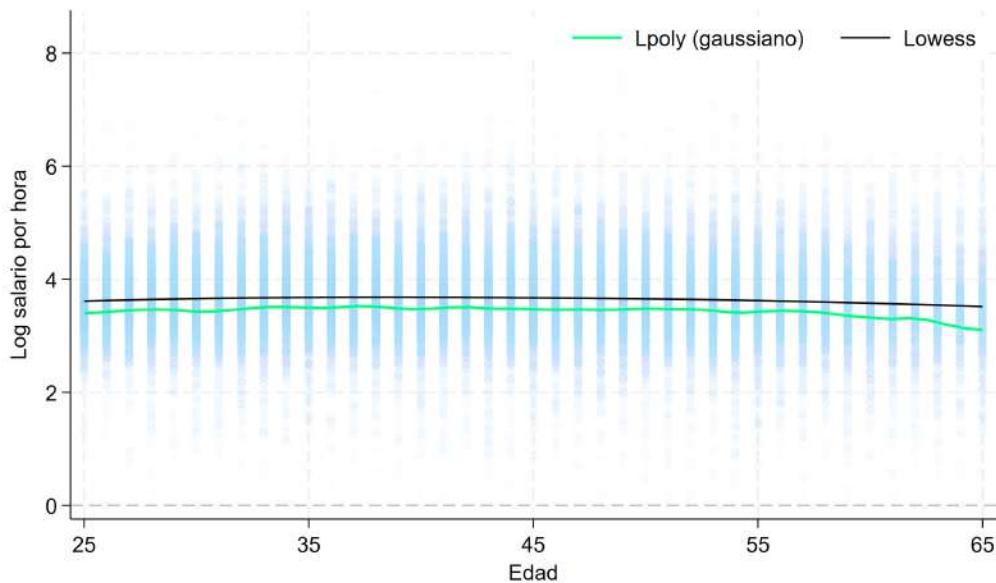
Fuente: Elaboración propia con datos de la ENOE. Proceso disponible en:  danifuentesga

Figura 102: Relación no paramétrica entre esolaridad y logaritmo del salario por hora en mujeres (ENOE 2018)



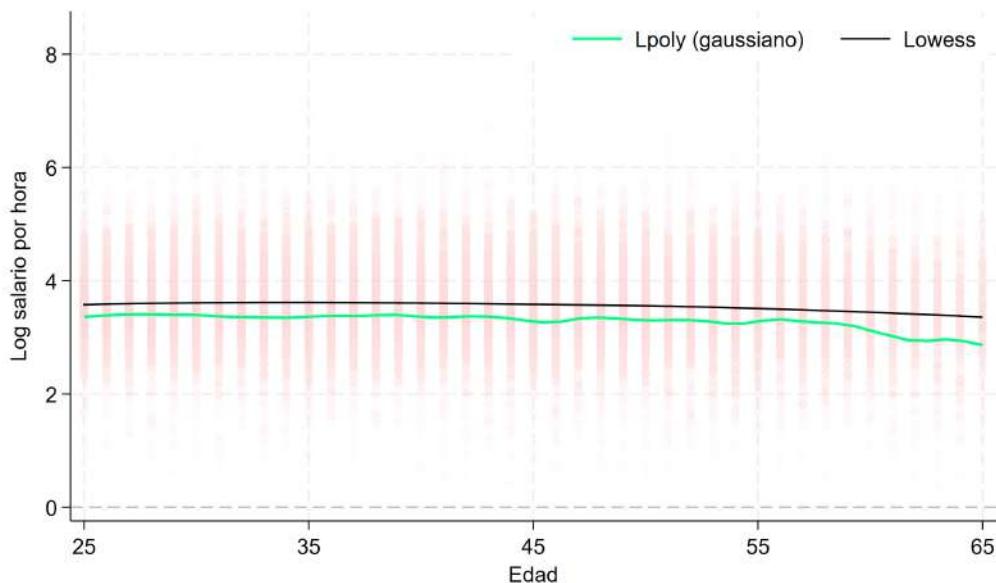
Fuente: Elaboración propia con datos de la ENOE. Proceso disponible en:  danifuentesga

Figura 103: Relación no paramétrica entre edad y logaritmo del salario por hora en hombres (ENOE 2024)



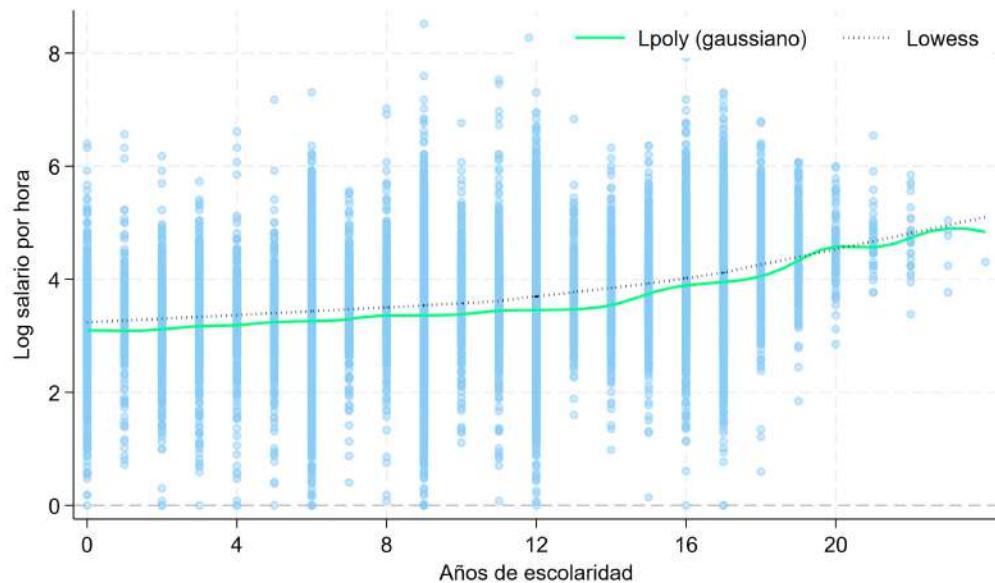
Fuente: Elaboración propia con datos de la ENOE. Proceso disponible en:  danifuentesga

Figura 104: Relación no paramétrica entre edad y logaritmo del salario por hora en mujeres (ENOE 2024)



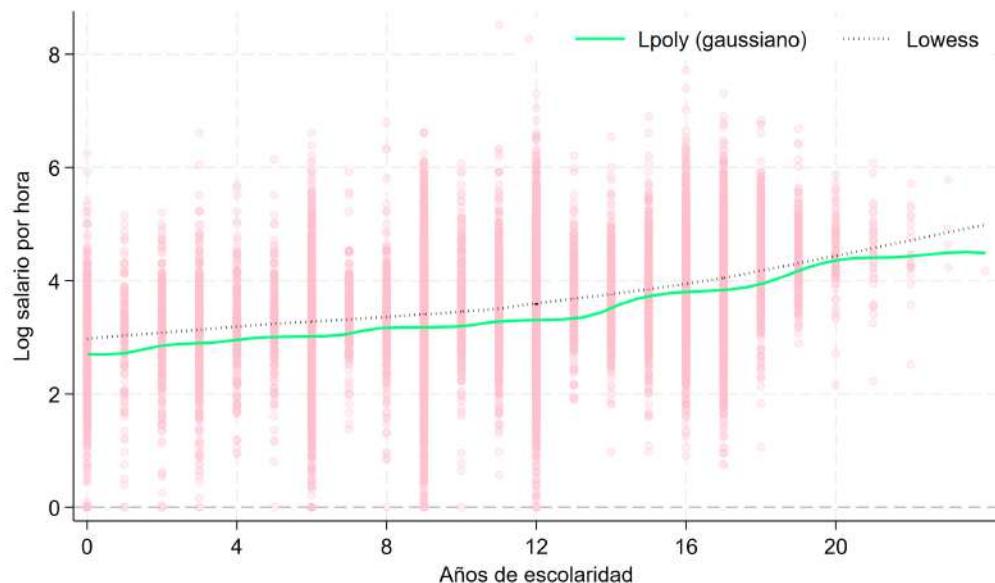
Fuente: Elaboración propia con datos de la ENOE. Proceso disponible en:  danifuentesga

Figura 105: Relación no paramétrica entre escolaridad y logaritmo del salario por hora en hombres (ENOE 2024)



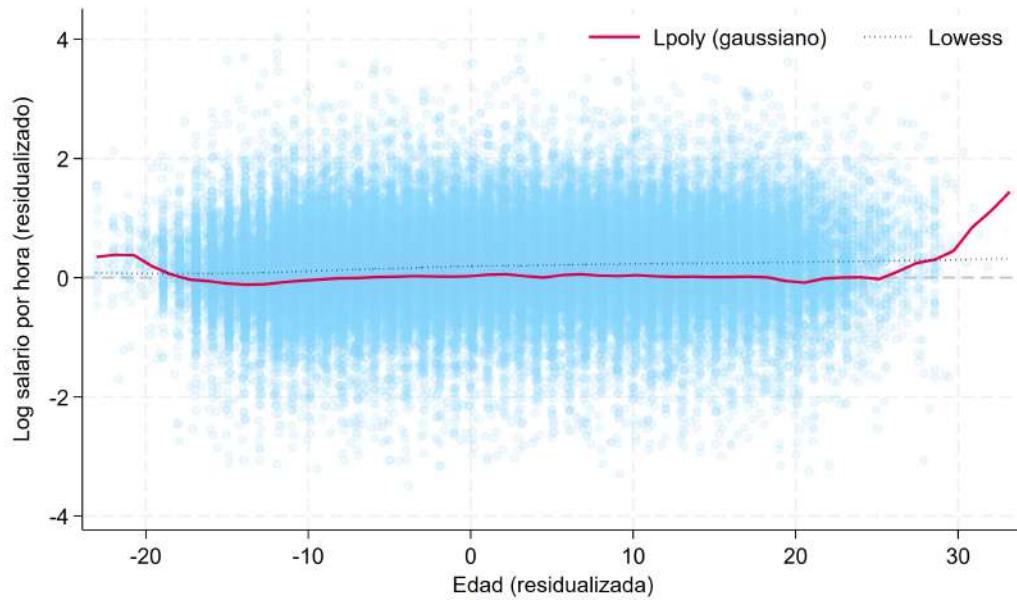
Fuente: Elaboración propia con datos de la ENOE. Proceso disponible en:  danifuentesga

Figura 106: Relación no paramétrica entre escolaridad y logaritmo del salario por hora en mujeres (ENOE 2024)



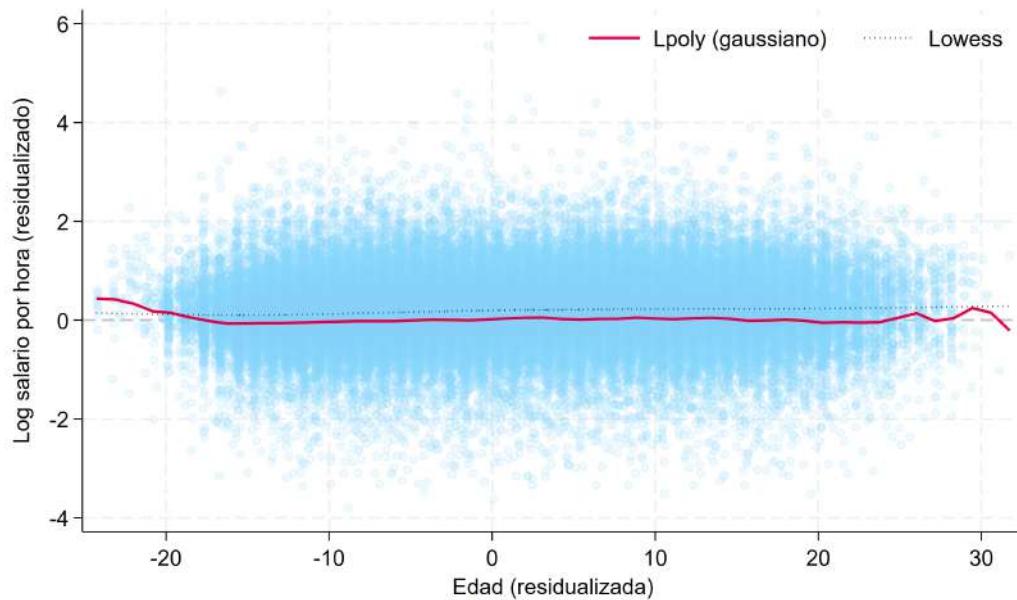
Fuente: Elaboración propia con datos de la ENOE. Proceso disponible en:  danifuentesga

Figura 107: Relación neta entre edad y logaritmo del salario por hora, controlando por escolaridad, sexo y ruralidad (ENOE,2018)



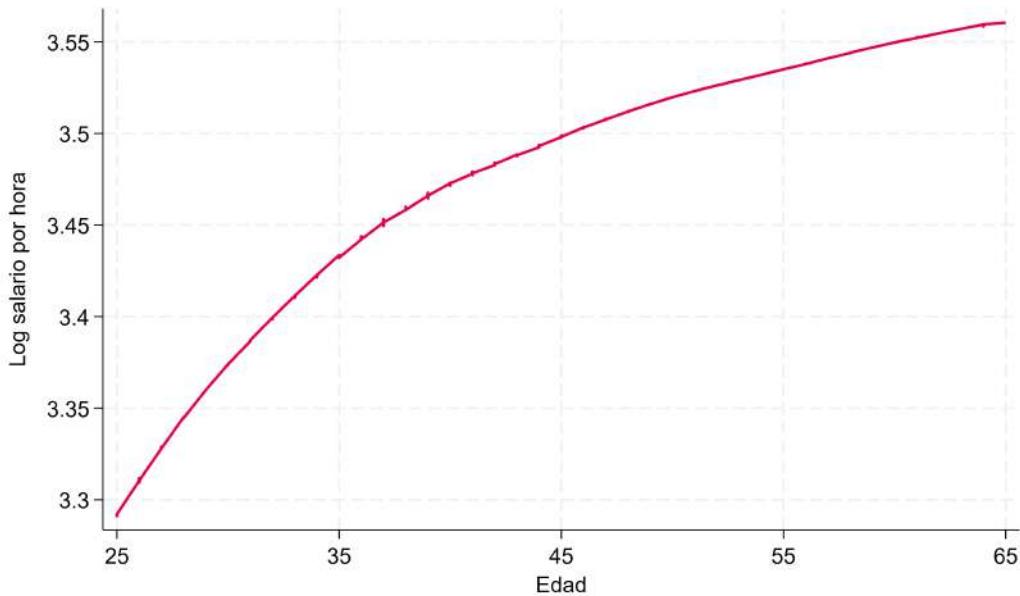
Fuente: Elaboración propia con datos de la ENOE. Proceso disponible en:  danifuentesga

Figura 108: Relación neta entre edad y logaritmo del salario por hora, controlando por escolaridad, sexo y ruralidad (ENOE,2022)



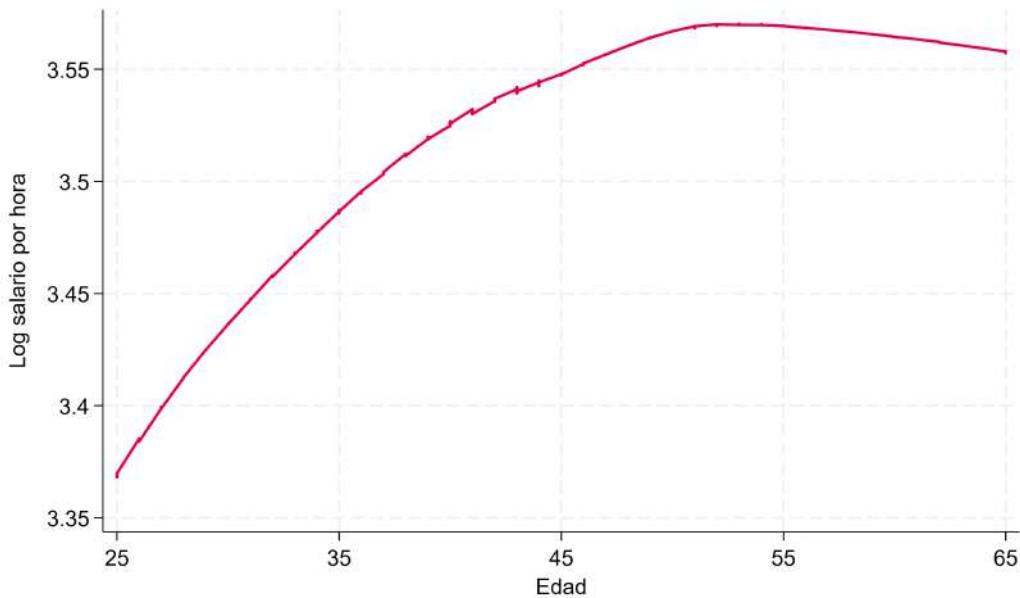
Fuente: Elaboración propia con datos de la ENOE. Proceso disponible en:  danifuentesga

Figura 109: Relación semiparamétrica entre edad e ingreso laboral horario en logaritmos (ENOE, 2018)



Fuente: Elaboración propia con datos de la ENOE. Proceso disponible en:  danifuentesga

Figura 110: Relación semiparamétrica entre edad e ingreso laboral horario en logaritmos (ENOE, 2022)



Fuente: Elaboración propia con datos de la ENOE. Proceso disponible en:  danifuentesga

Los ejercicios de **regresión no paramétrica** muestran que la **ENOE** y la **ENIGH** son **comparables**, pues en ambos casos las gráficas de **densidad**, **kdensity**, **lowess** y **lpoly** revelan **tendencias muy similares** en la relación entre **log salario** y **edad**, así como entre **log salario** y **escolaridad**. De manera consistente, la estimación **semiparamétrica** de log ingreso contra edad confirma el mismo patrón, lo que respalda la validez de usar ambas encuestas de forma conjunta para analizar la evolución salarial.

## Problema 6: Imputación

---

### 6.1

Lee el artículo de **Campos Vázquez (2013)** en la revista *Ensayos de Economía* de la **UANL** y el apéndice del artículo de **Levy y López Calva (2016)** (<https://publications.iadb.org/handle/11319/7454>).

Resume rápidamente el primer artículo, los métodos de imputación utilizados, y explica un método de imputación adicional que *no sea explicado en el artículo (los artículos de Rubin en la literatura o su libro te pudiera ayudar)*.

#### **RESPUESTA:**

El ensayo analiza la evolución del mercado laboral en México con un enfoque en la calidad de la información estadística. Se subraya que una proporción creciente de trabajadores no declara ingresos, y que este grupo tiene características distintas al resto, lo cual genera sesgos en las estimaciones de salarios, pobreza y desigualdad.

Ante este problema, el autor argumenta la importancia de la **imputación de ingresos faltantes**. Se revisan métodos tradicionales que resultan inadecuados porque alteran la distribución real del ingreso y reducen su varianza. En contraste, se proponen técnicas más flexibles y no paramétricas, capaces de producir estimaciones más realistas y útiles para la investigación y el diseño de políticas públicas.

El trabajo concluye que, sin un tratamiento adecuado de los datos faltantes, las mediciones del bienestar y de la desigualdad en México pueden estar distorsionadas. Por ello, la imputación no es solo un ajuste estadístico, sino una herramienta central para interpretar correctamente la dinámica laboral y orientar decisiones de política económica y social.

Un método alternativo de imputación es la **imputación múltiple** propuesta por Rubin. La idea central es no imputar un único valor faltante, sino generar **varias versiones completas de la base de datos**, donde los valores imputados se obtienen a partir de un

modelo estadístico que incorpora la incertidumbre de manera explícita.

El procedimiento básico es el siguiente:

1. Estimar un modelo predictivo de los datos faltantes utilizando las variables observadas.
2. Generar varias imputaciones distintas (por ejemplo, 5 o 10), de modo que cada base tenga valores imputados diferentes que reflejan la variabilidad inherente al proceso.
3. Analizar cada base por separado y combinar los resultados finales usando reglas de agregación que promedian estimadores y ajustan la varianza para reflejar la incertidumbre por los datos faltantes.

La ventaja principal es que, a diferencia de los métodos determinísticos, la **imputación múltiple conserva la varianza** y ofrece intervalos de confianza más realistas. Este método es especialmente útil cuando los datos faltantes son numerosos o dependen de múltiples factores, ya que evita el sesgo de imputar siempre un único valor “promedio” y permite análisis estadísticamente válidos

## 6.2

Para la **ENOE 2005–2025**, ¿cómo cambia el % de **trabajadores que no reportan ingresos** (pero sí son trabajadores)?

Realiza análisis por grupo: **formales vs informales, grupo de educación, sexo**, etc.

**¿El cambio en el no reporte podrías considerarlo aleatorio? Explica.**

### **RESPUESTA:**

Al ver los resultados , el cambio en el no reporte **NO** puede ser considerado como **aleatorio**, pues se observa una **tendencia creciente** y persistente en el tiempo (2005–2025). En el caso del **sexo**, las dos series se mueven casi en **paralelo** y la brecha es pequeña, lo que sugiere diferencias cercanas al **azar**. Sin embargo, al considerar **educación, formalidad y ruralidad** se aprecian patrones sistemáticos. Por ejemplo, los trabajadores del **sector informal** resportan sistemáticamente menos que los trabajadores **formales**. Los trabajadores con

mayores años de escolaridad reportan sistemáticamente menos que los de menor escolaridad, i.e los mas *educados* reportan **menos** sus ingresos. Asimismo, trabajadores en el sector **rural** reportan sistemáticamente menos que los trabajadores del sector **urbano**.

**¿Cuántos o qué % no contestan la pregunta de ingreso pero sí contestan la pregunta de rangos de salario mínimo (SM)?**

**RESPUESTA:**

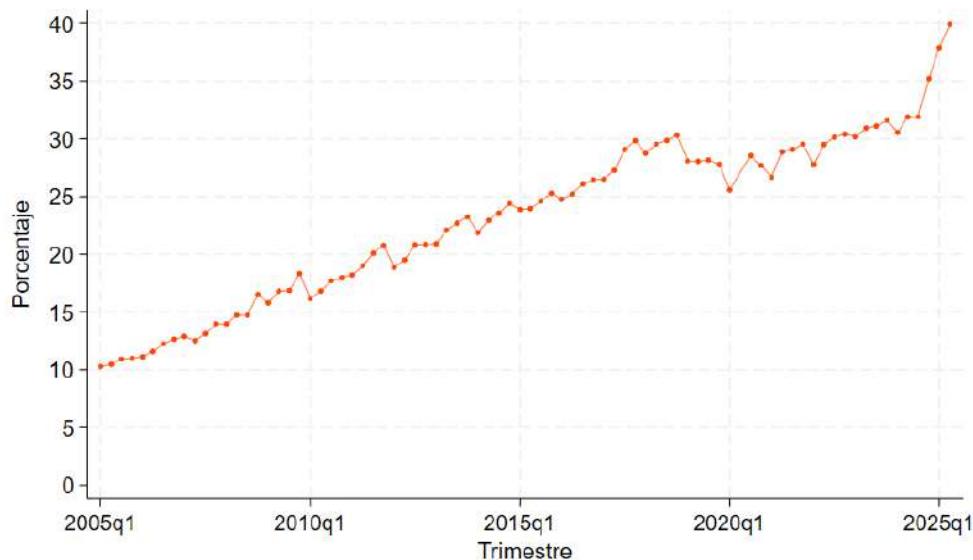
De acuerdo con la Figura 111 , el porcentaje de trabajadores que no responden la **pregunta directa de ingresos** pasó de alrededor del 10% en 2005 a cerca del 40% en 2025. Dentro de este grupo, una fracción relevante declara únicamente en **rangos de salario mínimo**, lo que representa entre 5% y 15% de los ocupados según el trimestre. En promedio, esto implica que cerca de **un tercio a la mitad** de quienes no reportan ingreso utilizan la opción de rangos salariales.

**¿Cómo ha cambiado ese % en el tiempo del total que no contesta la pregunta?**

**RESPUESTA:**

Entre 2005 y 2025, la fracción que responde sólo en **rangos salariales** dentro de los que no declaran ingreso se ha mantenido **estable**, alrededor de un **tercio a la mitad** del total.

Figura 111: Evolución trimestral del porcentaje de trabajadores que no reportan ingreso (ENOE, 2005-2025q2)



Fuente: Elaboración propia con datos de la ENOE. Proceso disponible en:  danifuentesga

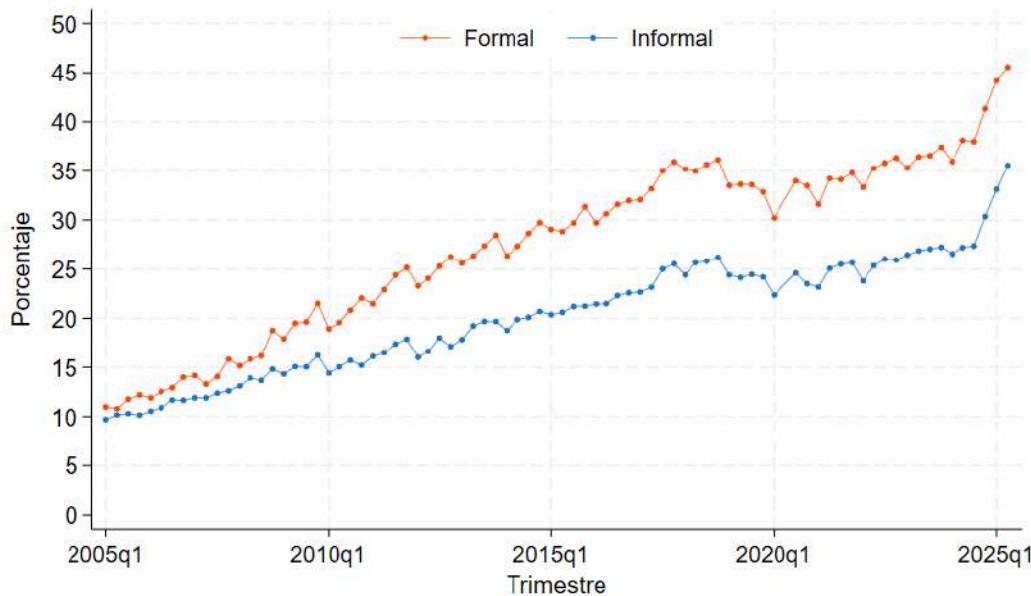
Figura 112: Evolución trimestral del porcentaje de trabajadores que no responde ingresos y rango de salario (ENOE, 2005-2025q2)



Nota: Porcentaje de trabajadores que no respondieron la pregunta de ingresos o que sólo declararon en rango salarial.

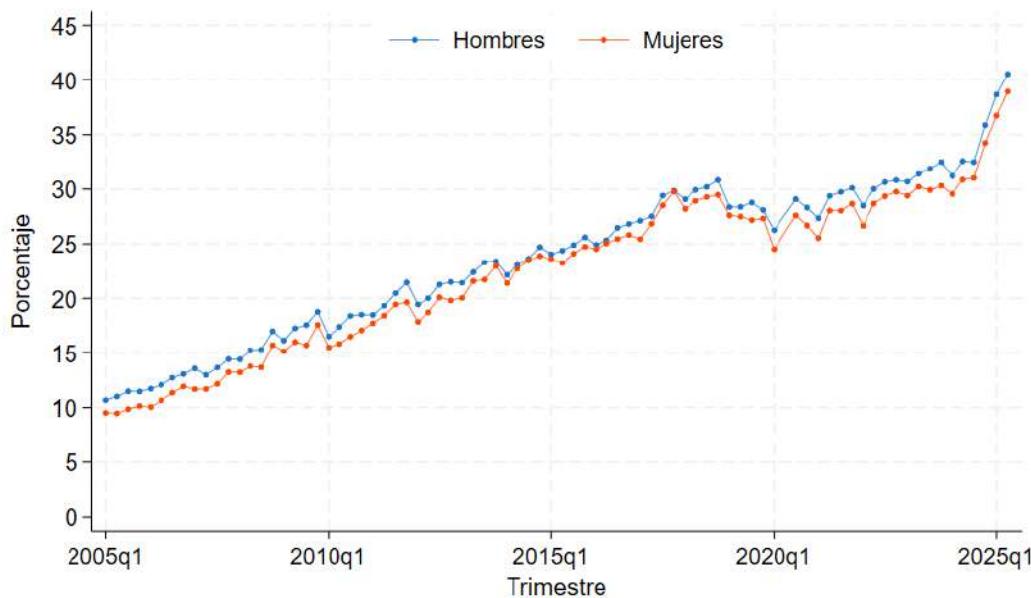
Fuente: Elaboración propia con datos de la ENOE. Proceso disponible en:  danifuentesga

Figura 113: Evolución trimestral del porcentaje de trabajadores que no reportan ingreso segun situación formal (ENOE, 2005-2025q2)



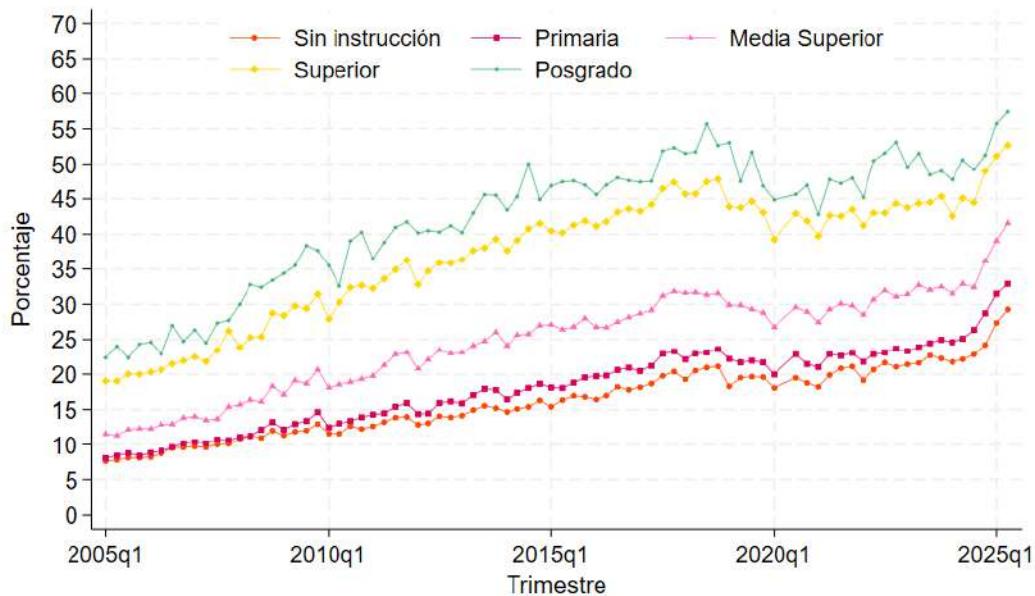
Fuente: Elaboración propia con datos de la ENOE. Proceso disponible en:  danifuentesga

Figura 114: Evolución trimestral del porcentaje de trabajadores que no reportan ingreso por sexo (ENOE, 2005-2025q2)



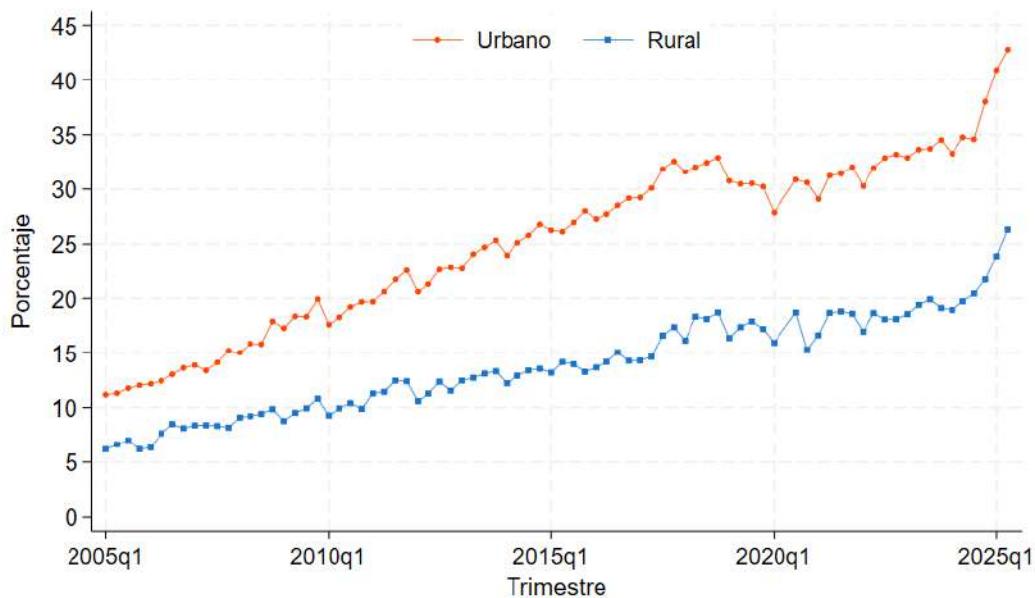
Fuente: Elaboración propia con datos de la ENOE. Proceso disponible en:  danifuentesga

Figura 115: Evolución trimestral del porcentaje de trabajadores que no reportan ingreso por nivel educativo (ENOE, 2005-2025q2)



Fuente: Elaboración propia con datos de la ENOE. Proceso disponible en:  danifuentesga

Figura 116: Evolución trimestral del porcentaje de trabajadores que no reportan ingreso por sector (ENOE, 2005-2025q2)



Fuente: Elaboración propia con datos de la ENOE. Proceso disponible en:  danifuentesga

## 6.3

Replica el cálculo de **pobreza laboral de CONEVAL** con la **ENOE 2005–2025**.

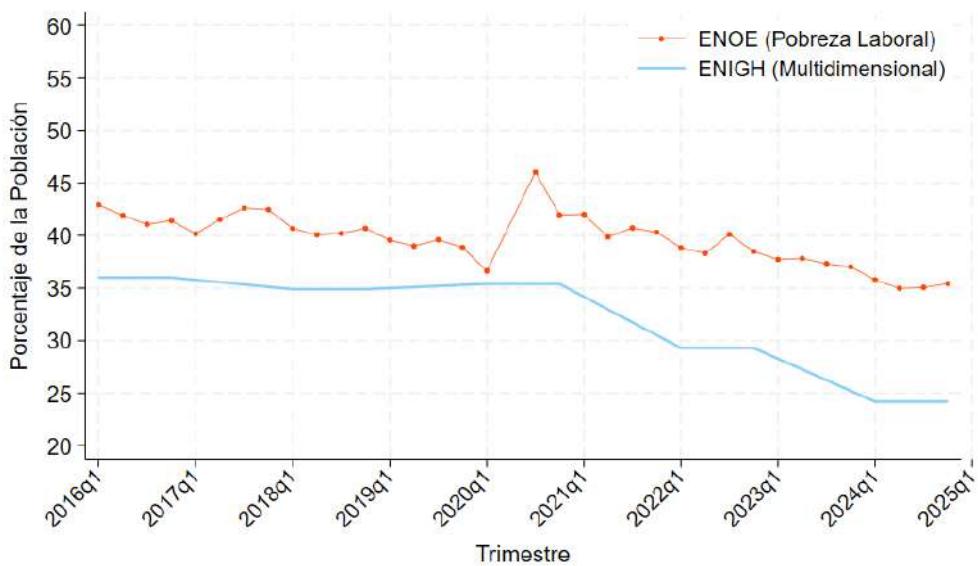
Los programas están disponibles en la red. Asegúrate de entender los *supuestos que realizan para los que no declaran ingresos*.

Descarga la **serie de tiempo de pobreza multidimensional** usando **ENIGH** para los últimos **10 años**.

Realiza dos figuras y dos mapas:

- Una figura en **serie de tiempo** con dos líneas de pobreza: **ENOE** y **ENIGH**,
- Otra figura tipo **scatter plot** con **ENOE** en el eje x y **ENIGH** en el eje y, contrastando la **correlación entre ambas**,
- Los **mapas** deben ser hechos en **R** para calcular la pobreza a **nivel entidad federativa** para el **último año**.

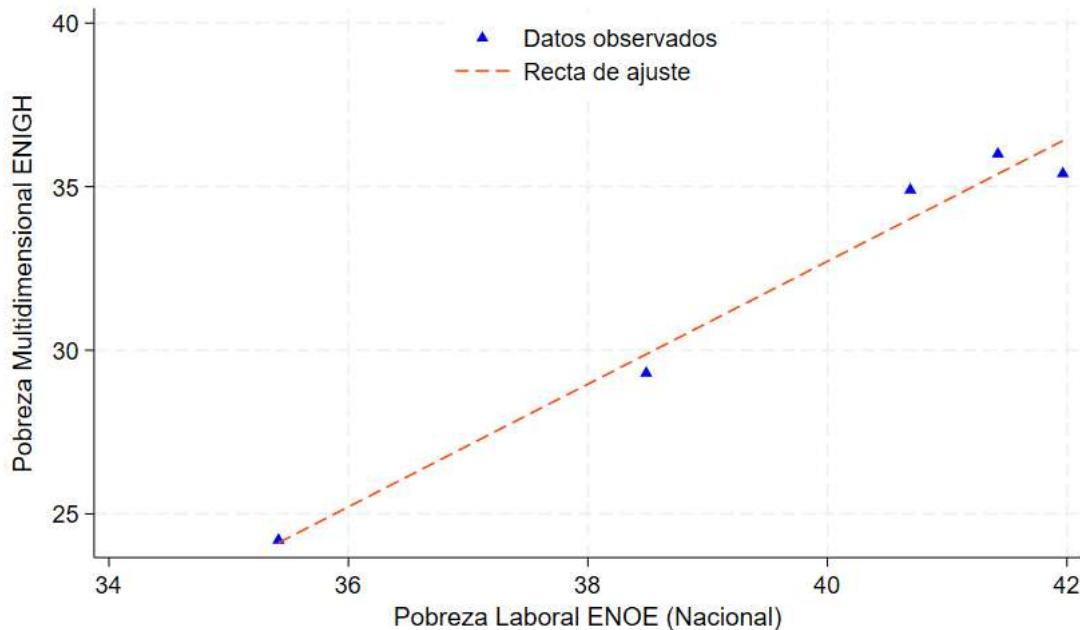
Figura 117: Evolución De Poblacion en Pobreza Multidimensional y Pobreza Laboral,(ENIGH y ENOE, 2016-2024)



*Nota:* Periodo 2020q2 no se levantó debido a la pandemia, por lo que la configuración del programa üne "la serie a pesar de haber un trimestre faltante, por lo que no queda exactamente igual a la serie de CONEVAL la cual presenta un "salto en este periodo.

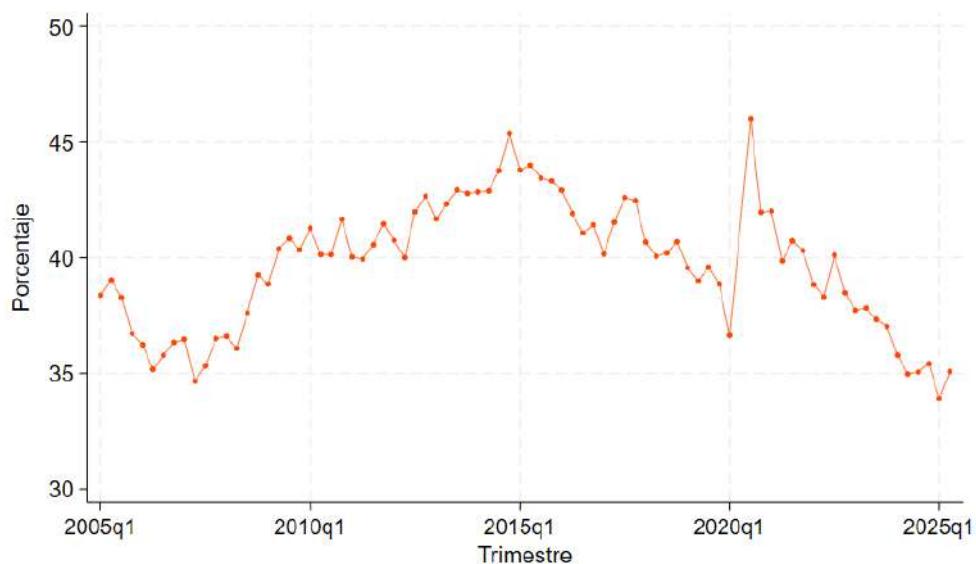
Fuente: Elaboración propia con datos de la ENOE e INEGI. Proceso disponible en:  danifuentesga

Figura 118: Relación entre Pobreza Multidimensional y Pobreza Laboral,(ENIGH y ENOE, 2016-2024)



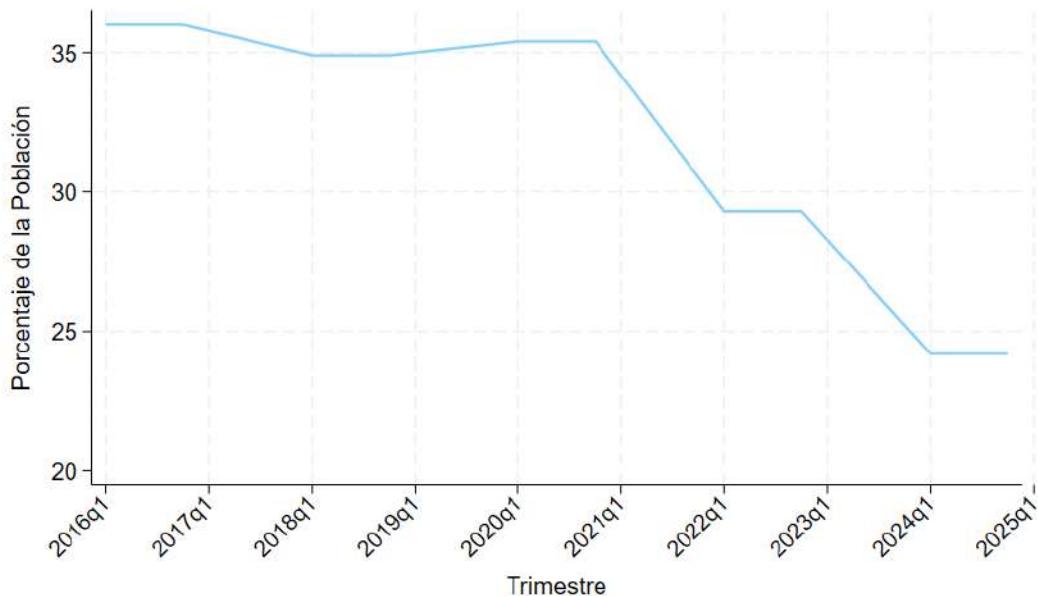
Fuente: Elaboración propia con datos de la ENOE e INEGI. Proceso disponible en:  danifuentesga

Figura 119: Evolución De Poblacion en Pobreza Laboral (ENOE, 2005-2025q2)



Fuente: Elaboración propia con datos de la ENOE. Proceso disponible en:  danifuentesga

Figura 120: Evolución De Poblacion en Pobreza Multidimensional, INEGI (ENIGH, 2016-2024)



Fuente: Elaboración propia con datos del INEGI. Proceso disponible en:  danifuentesga

Figura 121: Porcentaje de la Poblacion en Pobreza Laboral,(ENOE, 2025 TRIMSTRE II)



Fuente: Elaboración propia con datos de la ENOE. Proceso disponible en:  danifuentesga

Figura 122: Porcentaje de la Población en Pobreza Multidimensional,(ENIGH, 2024)



Fuente: Elaboración propia con datos del INEGI. Proceso disponible en:  danifuentesga

## 6.4

Replica el artículo de Campos Vázquez (2013) para el periodo **2005–2025**, utilizando todos los trimestres.

Utilizarás **3 métodos de imputación**:

1. **Hotdeck**,
2. **Medias de grupos específicos** (*si este método no te gusta, puedes sustituirlo por uno de tu preferencia*).

En **dos gráficas por separado** incluirás lo siguiente:

- a) % de **Pobreza laboral** de acuerdo con cada método (*3 líneas en la figura: la oficial y los 2 métodos adicionales*),
- b) **Ingreso laboral per cápita** (*3 líneas en la figura: oficial y los 2 métodos*).

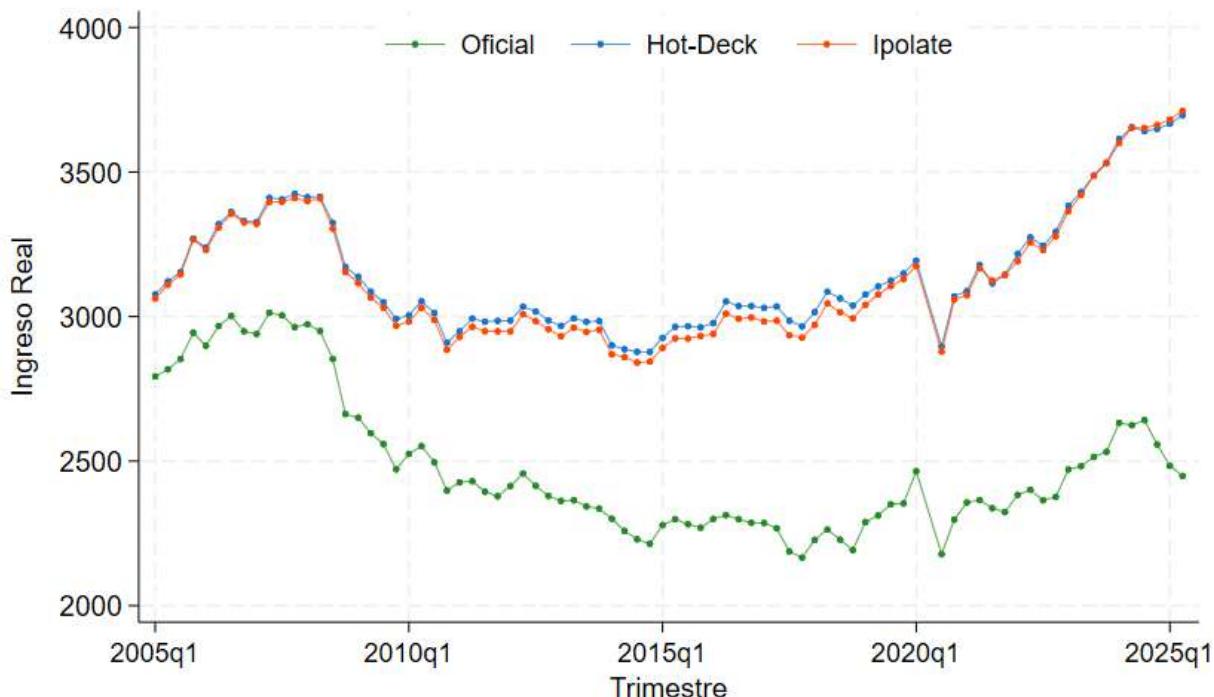
Gráficas que no se entiendan o no se vean bonitas serán castigadas.

Para este problema tendrás que tener acceso a las **líneas de bienestar** y a los **programas de cálculo del ITLP** por parte de **CONEVAL**, los cuales están disponibles en línea.

El CONEVAL realiza un **ajuste al ingreso**, asegúrate de que entiendes el *procedimiento utilizado* y que lo puedas mencionar.

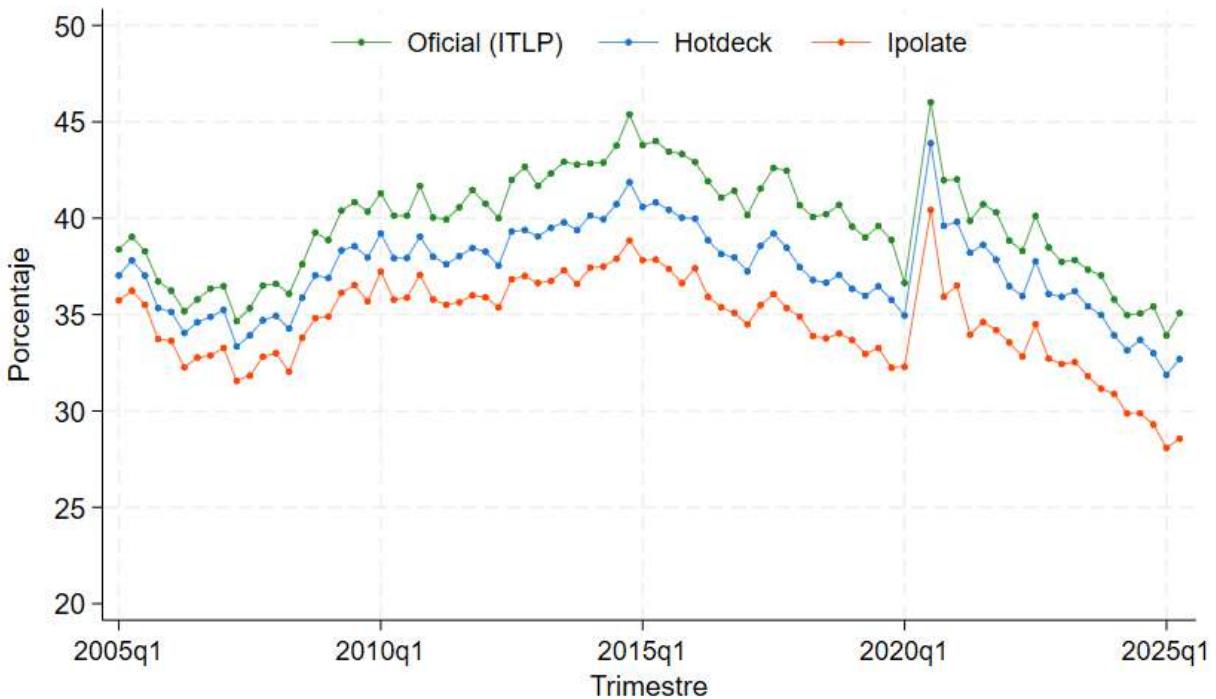
### **RESPUESTA:**

Figura 123: Ingreso Reportado vs Hotdeck vs iPolate (ENOE,2005-2025q2)



Fuente: Elaboración propia con datos de la ENOE. Proceso disponible en:  danifuentesga

Figura 124: % de la Población en ITLP: Reportado vs Hotdeck vs iPolate (ENOE,2005-2025q2)



Fuente: Elaboración propia con datos de la ENOE. Proceso disponible en:  danifuentesga

## 6.5

**¿Qué puedes concluir del problema de los ingresos faltantes y de la comparación entre métodos?**

También analiza la comparación entre ENIGH y ENOE.

### RESPUESTA:

Las gráficas muestran que el **problema de los ingresos faltantes en la ENOE** ha crecido de forma sostenida, llegando a más del 35 % de los trabajadores en 2025 (Figura 111). Esto afecta el cálculo de indicadores como el ingreso promedio y la pobreza laboral. Los métodos de imputación **Hot-Deck e Ipolate** (Figuras 123 y 124) corrigen parcialmente este sesgo.

Las **Figuras 117 y 118** muestran que la **ENOE (pobreza laboral)** y la **ENIGH**

(pobreza multidimensional) presentan niveles distintos, pero siguen **tendencias consistentes de reducción** en el periodo 2016–2024. La relación entre ambas medidas es fuertemente positiva: cuando baja la pobreza laboral en la ENOE, también disminuye la pobreza multidimensional en la ENIGH. Esto confirma que, aunque los indicadores capturan dimensiones distintas, **pueden usarse de forma complementaria** para analizar la evolución de la pobreza en México.

## 6.6

**¿Crees que los supuestos del método de imputación se cumplan? Explica y argumenta tu respuesta.**

### **RESPUESTA:**

La validez del **método de imputación** depende de que los ingresos faltantes sean **aleatorios condicionales** a las variables observadas (edad, escolaridad, ocupación, etc.). Esto se cumple, por ejemplo, si en la ENOE algunos encuestados olvidan decir su ingreso pero comparten características similares a quienes sí lo reportan, lo que permite imputar de manera confiable. Sin embargo, el supuesto no se cumple si quienes **no reportan** pertenecen de forma sistemática a un grupo distinto (por ejemplo, trabajadores con salarios muy altos que prefieren no contestar), ya que la imputación subestimaría esos ingresos. En resumen, el método funciona bien cuando el **no-respuesta es aleatoria**, pero puede sesgar los resultados si está ligada a **niveles específicos de ingreso**.

## 6.7

**La pobreza laboral aumentó con COVID y luego disminuye en 2022–2025. ¿A qué se debe?**

A nivel **entidad federativa**, realiza un *scatter* entre la **disminución de pobreza laboral** y los **cambios en empleo y salario** con datos del **IMSS** y la **ENOE**.

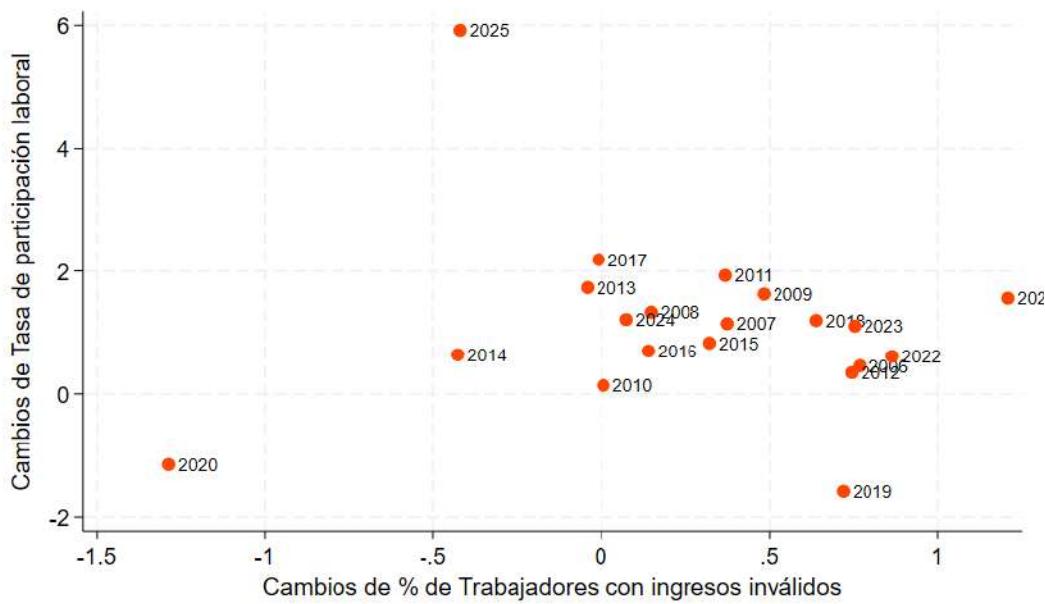
Incluye también los **cambios en participación laboral** y los **cambios en la proporción que no declaran ingresos** según la ENOE.

¿Qué podemos decir al respecto?

**RESPUESTA:**

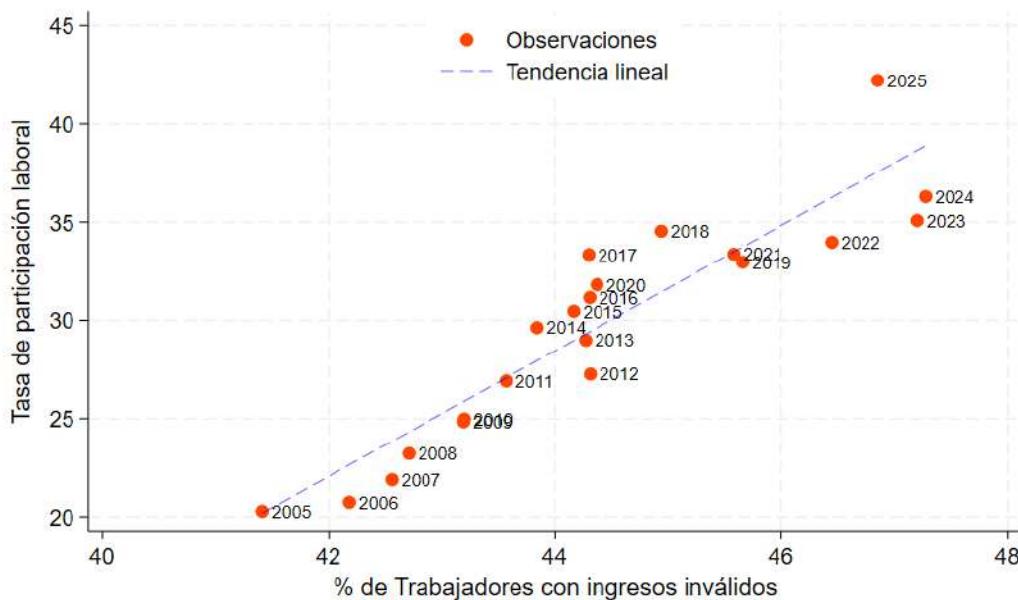
El aumento de la **pobreza laboral durante la COVID-19** se explica por la caída abrupta del empleo y del ingreso promedio (Figuras 127 y 128), así como por un repunte en la proporción de trabajadores que no declararon ingreso en la ENOE (Figuras 125 y 126). Posteriormente, entre 2022 y 2025, la recuperación del empleo formal y la mejora en salarios redujeron la pobreza laboral, aunque de manera heterogénea entre entidades federativas. En suma, la evidencia sugiere que la **dinámica de la pobreza laboral está estrechamente ligada a choques en empleo, salarios y calidad del reporte de ingresos**.

Figura 125: Correlación entre cambios en trabajadores con ingresos inválidos y cambios la tasa de participación laboral (ENOE,2005-2025q2)



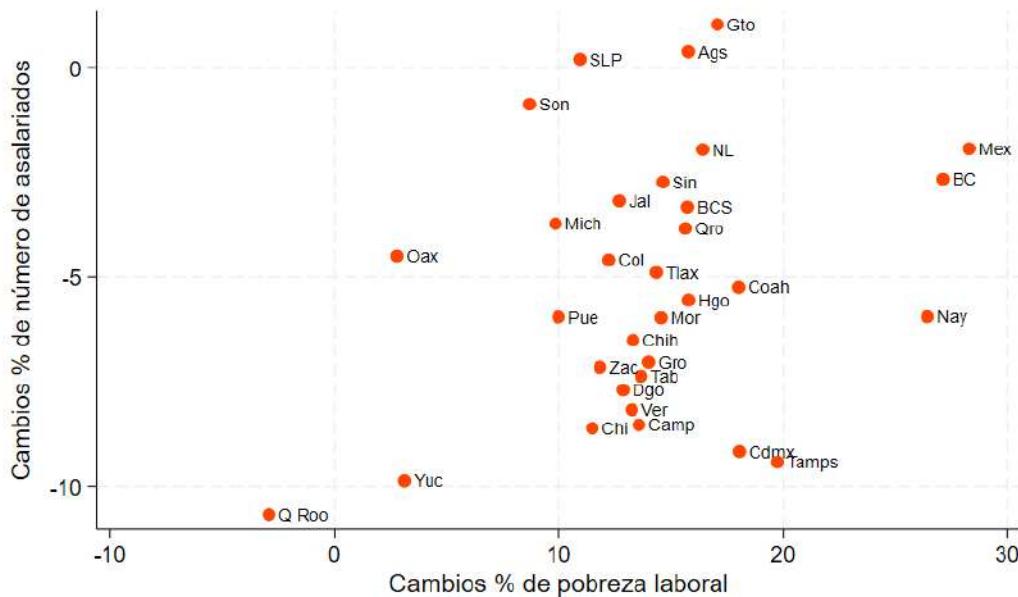
Fuente: Elaboración propia con datos de la ENOE. Proceso disponible en:  danifuentesga

Figura 126: Correlación entre trabajadores con ingresos inválidos y la tasa de participación laboral (ENOE,2005-2025q2)



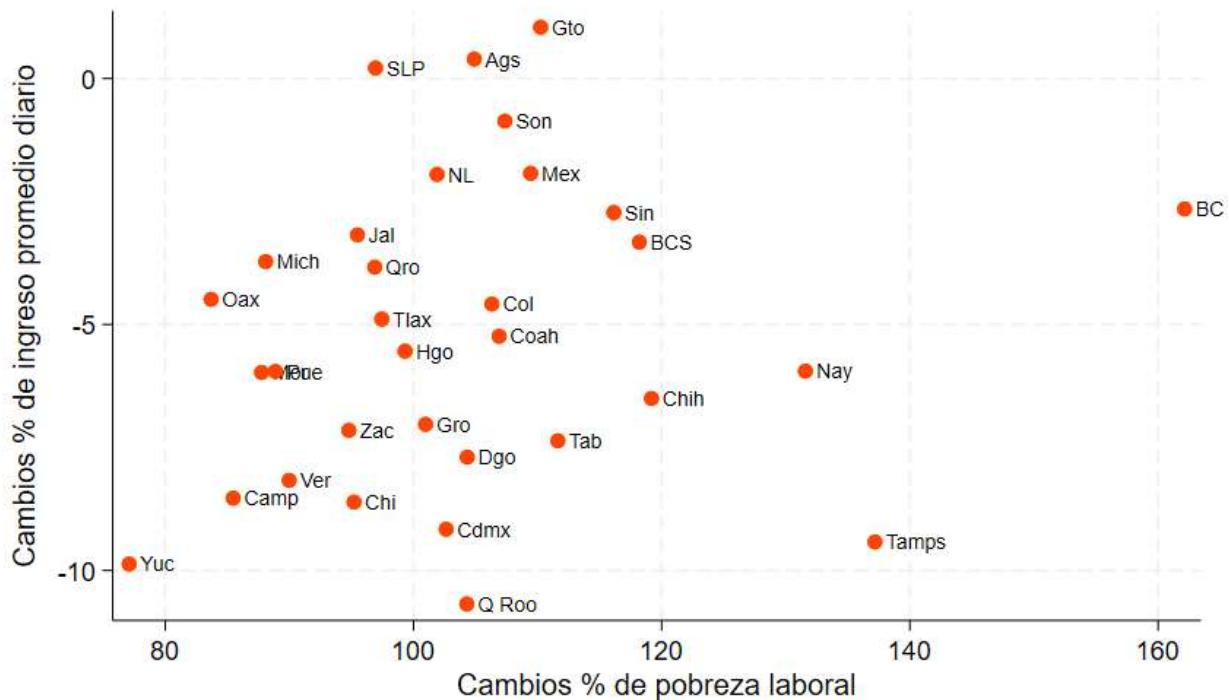
Fuente: Elaboración propia con datos de la ENOE. Proceso disponible en:  danifuentesga

Figura 127: Relación de cambios porcentuales entre de pobreza laboral y número de asalariados (2018q2,2025q2)



Fuente: Elaboración propia con datos de la ENOE e IMSS. Proceso disponible en:  danifuentesga

Figura 128: Relación de cambios porcentuales entre de pobreza laboral e ingreso promedio diario (2018q2,2025q2)



Fuente: Elaboración propia con datos de la ENOE e IMSS. Proceso disponible en:  danifuentesga

## Apéndice

---

### A : Regresión Cuantil

#### Cuantil poblacional incondicional

$$q = \Pr(y \leq \mu_q) = F_y(\mu_q)$$

$$\mu_q = F_y^{-1}(q)$$

- $y$ : variable aleatoria (por ejemplo, salario),
- $\mu_q$ : cuantil incondicional de orden  $q$ ,
- $F_y(\cdot)$ : función de distribución acumulada de  $y$ ,
- $F_y^{-1}(q)$ : función cuantil inversa.

El salario  $\mu_q$  es tal que el  $q$ . En el contexto de salarios, si  $q = 0,75$ , entonces el 75 % de la población gana menos o igual que  $\mu_{0,75}$ .

#### Cuantil condicional

$$Q_\tau(y | x) = F_{y|x}^{-1}(\tau) = x^\top \beta_\tau$$

$$\Pr(y \leq Q_\tau(y | x) | x) = \tau$$

- $x \in \mathbb{R}^k$ : vector de variables explicativas (e.g. educación),
- $\beta_\tau \in \mathbb{R}^k$ : coeficientes para el cuantil  $\tau$ ,
- $Q_\tau(y | x)$ : cuantil- $\tau$  condicional de  $y$  dado  $x$ ,
- $F_{y|x}$ : función de distribución condicional de  $y$  dado  $x$ .

Esta expresión estima cuánto gana una persona que se encuentra en el percentil  $\tau$ , dado

su nivel de educación.

### Función de pérdida asimétrica

$$Q_N(\beta_\tau) = \sum_{i:y_i \geq x_i^\top \beta_\tau} \tau |y_i - x_i^\top \beta_\tau| + \sum_{i:y_i < x_i^\top \beta_\tau} (1 - \tau) |y_i - x_i^\top \beta_\tau|$$

- $y_i$ : salario observado del individuo  $i$ ,
- $x_i$ : vector de características del individuo  $i$ ,
- $\beta_\tau$ : parámetro cuantil a estimar,
- $N$ : número de observaciones.

Este criterio penaliza de forma desigual los errores por exceso o defecto. Por ejemplo, al estimar el percentil 90, subestimar salarios se penaliza más que sobreestimarlos.

### Distribución asintótica del estimador

$$\sqrt{N}(\hat{\beta}_\tau - \beta_\tau) \xrightarrow{d} \mathcal{N}(0, A^{-1}BA^{-1})$$

$$A = \frac{1}{N} \sum_{i=1}^N f_{u_\tau}(0 \mid x_i) x_i x_i^\top$$

$$B = \frac{1}{N} \sum_{i=1}^N \tau(1 - \tau) x_i x_i^\top$$

- $\hat{\beta}_\tau$ : estimador del cuantil  $\tau$ ,
- $f_{u_\tau}(0 \mid x_i)$ : densidad del residuo evaluada en cero,
- : convergencia en probabilidad.

Con muestras grandes, el estimador cuantil se comporta normalmente. Esto permite

construir intervalos de confianza para el efecto de la educación en, por ejemplo, el percentil 75 del salario.

### Modelo con heterocedasticidad multiplicativa

$$y = x^\top \beta + u$$

$$u = x^\top \alpha \cdot \varepsilon$$

$$\varepsilon \sim iid [0, \sigma^2]$$

$$x^\top \alpha > 0$$

- $\alpha$ : vector que gobierna la heterocedasticidad,
- $\varepsilon$ : ruido base homocedástico,
- $u$ : término de error general,
- $x$ : características individuales.

Este modelo permite que la varianza del salario dependa de la educación. La dispersión puede ser mayor entre personas con títulos superiores que entre quienes solo tienen educación básica.

### Cuantil condicional con heterocedasticidad

$$\mu_\tau(x, \beta, \alpha) = x^\top \beta + Q_\tau(\varepsilon) \cdot x^\top \alpha$$

El salario en el cuantil  $\tau$  es afectado tanto por el efecto promedio de la educación como por la dispersión que varía con el nivel educativo.

### Derivación desde definición de cuantil

$$\begin{aligned}\tau &= \Pr(y \leq \mu_\tau(x, \beta, \alpha)) \\ &= \Pr\left(\varepsilon \leq \frac{\mu_\tau(x, \beta, \alpha) - x^\top \beta}{x^\top \alpha}\right) \\ &= F_\varepsilon\left(\frac{\mu_\tau(x, \beta, \alpha) - x^\top \beta}{x^\top \alpha}\right)\end{aligned}$$

Se busca el nivel de salario tal que, para un nivel dado de educación, la probabilidad de ganar menos o igual sea exactamente  $\tau$ .

### Función cuantil explícita

$$\mu_\tau(x, \beta, \alpha) = x^\top (\beta + \alpha \cdot F_\varepsilon^{-1}(\tau))$$

Cada cuantil de salario tiene su propia pendiente respecto a la educación. Un año adicional puede impactar más a quienes están en la parte alta de la distribución salarial.

### Cuantil condicional vs. incondicional

$$Q_\tau(Y | X) = X\beta_\tau$$

$$Q_\tau(Y) \neq Q_\tau(X)\beta_\tau$$

Conocer el cuantil del salario por nivel educativo no implica conocer el cuantil del salario de la población total.

### Distribución marginal a partir de cuantiles condicionales

$$F_Y(y | X) = \int_0^1 \mathbb{1}[X\beta_\tau < y] d\tau$$
$$F_Y(y) = \int_0^1 \int \mathbb{1}[X\beta_\tau < y] d\tau dF_X(x)$$
$$Q_\tau(Y) = \inf \{y : F_Y(y) \geq \tau\}$$

La distribución total del salario se obtiene combinando las distribuciones condicionales de cada grupo educativo, ponderadas por su frecuencia en la población.

## B:Estimación no Paramétrica, Trade-off Varianza-Sesgo

$$\hat{m}_h(x) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{x - X_i}{h}\right) Y_i \quad \text{Estimador Kernel} \quad (1)$$

$$\mathbb{E}[\hat{m}_h(x)] = m(x) + \frac{1}{2}h^2 m''(x)\mu_2(K) + o(h^2) \quad \text{Sesgo de orden } h^2 \quad (2)$$

$$\text{Bias}[\hat{m}_h(x)] = \frac{1}{2}h^2 m''(x)\mu_2(K) \quad \text{Def. Sesgo} \quad (3)$$

$$\text{Var}[\hat{m}_h(x)] = \frac{1}{nh} \sigma^2(x) R(K) f^{-1}(x) + o\left(\frac{1}{nh}\right) \quad \text{Varianza de orden } \frac{1}{nh} \quad (4)$$

$$\text{MSE}(h) = \text{Bias}^2[\hat{m}_h(x)] + \text{Var}[\hat{m}_h(x)] \quad \text{Def. MSE} \quad (5)$$

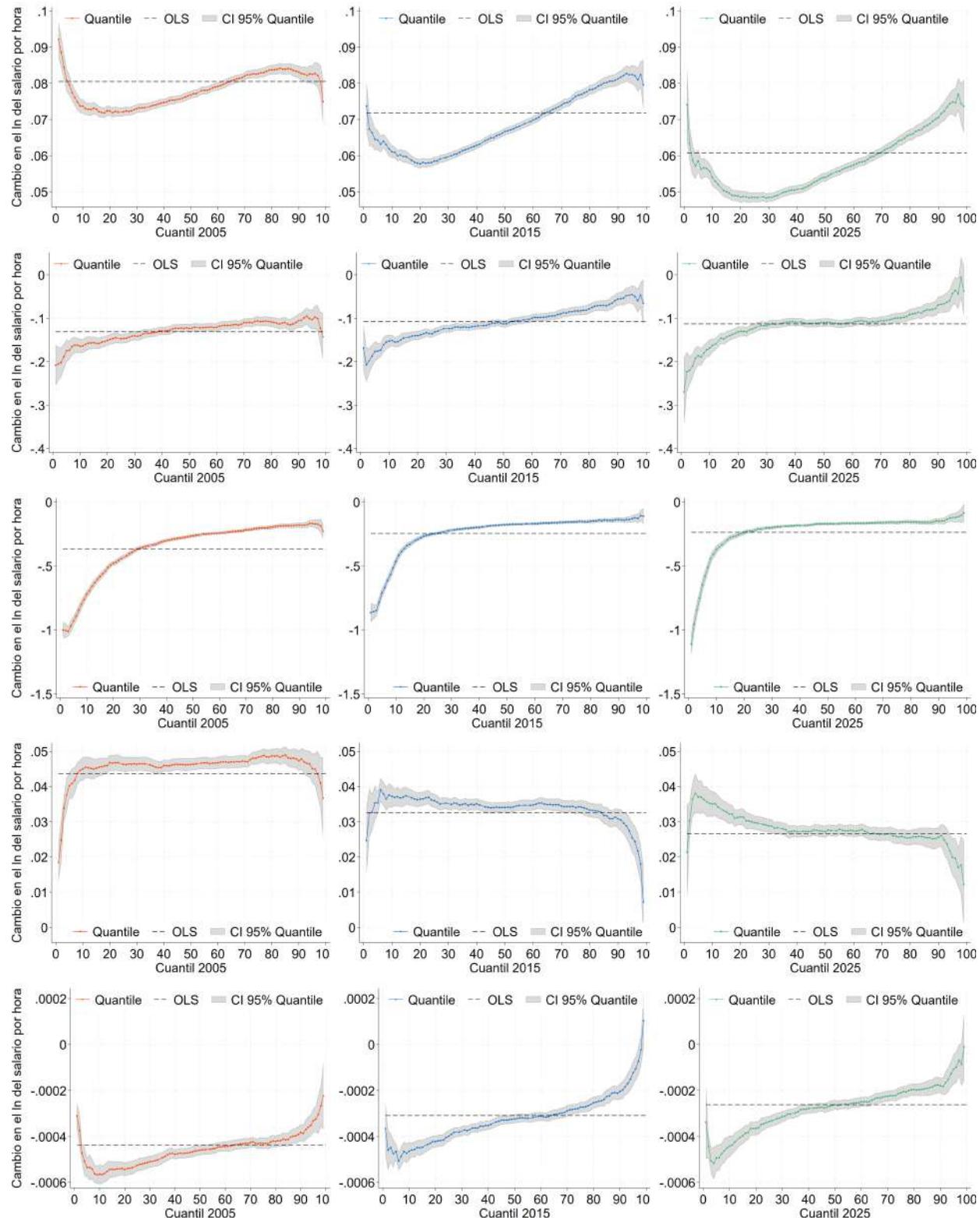
$$= \frac{1}{4}h^4 \left(m''(x)\mu_2(K)\right)^2 + \frac{1}{nh} \sigma^2(x) R(K) f^{-1}(x) \quad \text{Expresión Final} \quad (6)$$

$h \uparrow \Rightarrow \text{Var} \downarrow, \text{ Bias} \uparrow$       **Suaviza pero sesga**      (7)

$h \downarrow \Rightarrow \text{Var} \uparrow, \text{ Bias} \downarrow$       **Preciso pero ruidoso**      (8)

## C : Regresiones Cuantiles ENOE

Figura 129: Regresiones Cuantiles (ENOE: 2005,2015,2025)



Fuente: Elaboración propia con datos de la ENIGH. Proceso disponible en:  danifuentesga