

EDUCACIÓN EN MÉXICO

EL COLEGIO DE MÉXICO

SEPTIEMBRE, 2025

Andrea Fastlicht Perelman
andrea.fastlicht@banxico.org.mx
Banco de México

EDUCACIÓN COMO CONCEPTO

Capital humano

- La educación incrementa la productividad y las oportunidades de ingreso, con retornos privados de hasta 10% para ciertos contextos (Duflo, 2001; Angrist & Krueger, 1991).
- Los retornos sociales suelen ser menores, ya que incluyen subsidios públicos; aún así, generan beneficios colectivos de crecimiento y bienestar (Psacharopoulos & Patrinos, 2018).

Impacto en salud

- Un año adicional de escolaridad aumenta la esperanza de vida en 1.7 años (Lleras-Muney, 2005).
- También, la educación reduce la mortalidad al facilitar empleos menos riesgosos y mayor acceso a información (Cutler & Lleras-Muney, 2006).

Efectos en el desarrollo

- Más allá de mejorar las habilidades individuales, se requieren entornos institucionales y económicos adecuados para aprovechar la educación (ej. diferencias China vs. India: Banerjee & Duflo, 2005).

Oferta vs. demanda educativa

- No basta con construir escuelas; la asistencia depende de que padres e hij@s perciban y valoren los retornos de la educación (Banerjee & Duflo, 2011).

IMPORTANCIA DE LA EDUCACIÓN

Relevancia en México

- El rezago educativo es una barrera crítica para la movilidad social (Esquivel, 2024).
- Perpetúa desigualdades socioeconómicas e intergeneracionales (Ruhm, 2004).

Impactos

- Reduce oportunidades de desarrollo personal y académico (Reardon, 2011; Giorguli et al., 2010).
- Limita la inserción laboral y refuerza ciclos de pobreza (Giorguli et al., 2010).
- Resta potencial de crecimiento económico sostenible y capital humano (Gil-Lacruz et al., 2020).

Marco legal

- Derecho garantizado en el Art. 3 Constitucional.
- Reconocido como carencia social en la Ley General de Desarrollo Social.

Rezago educativo

Diferencia entre el nivel educativo esperado por edad vs. el nivel realmente alcanzado.

REZAGO EDUCATIVO EN MÉXICO

Tendencias nacionales

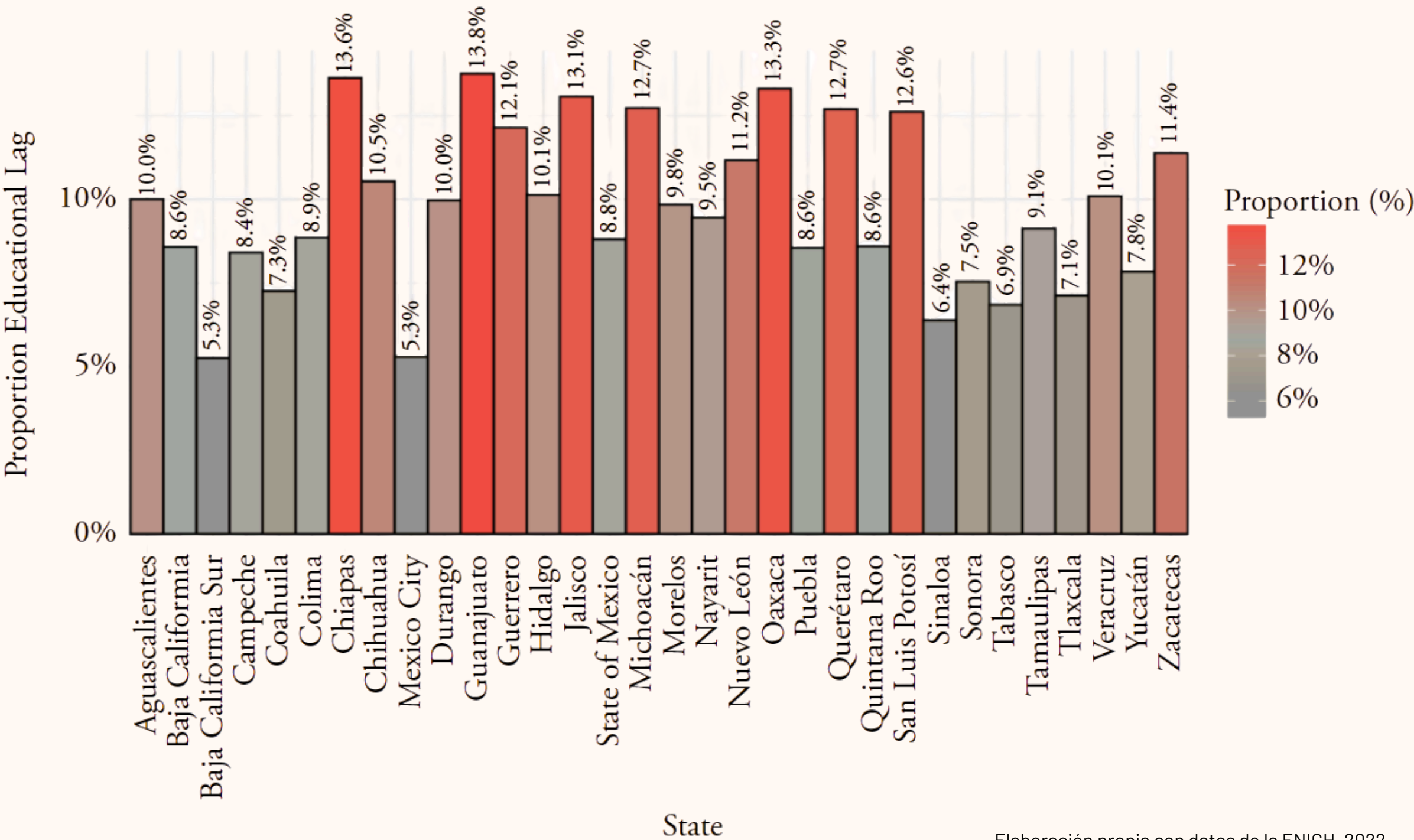
- Abandono en preparatoria:
 - 17.5% en 2000 → 8.7% en 2023 (INEGI, 2023).
- Abandono en secundaria:
 - 8.3% → 6.0%
 - Pero se mantiene estancado.

Grupos vulnerables

- Poblaciones rurales, indígenas y de bajos ingresos con mayor rezago (SEP, 2024).

La tasa de finalización de la primaria ya alcanza casi el 90% a nivel mundial (UNESCO, 2021).

Figure A3. Proportion of Individual's Educational Lag by State in Mexico



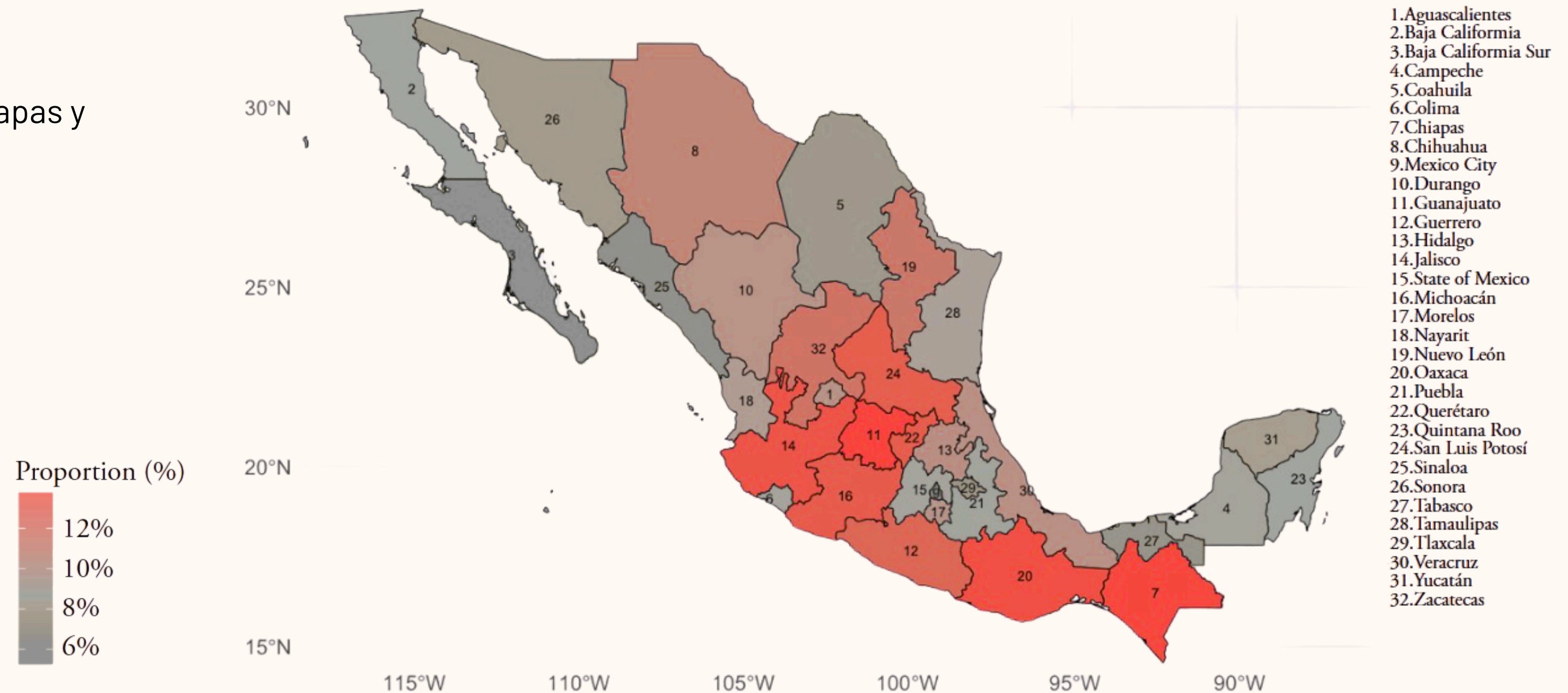
Elaboración propia con datos de la ENIGH, 2022.

REZAGO EDUCATIVO EN MÉXICO

Disparidades regionales

- Menor rezago: CDMX y BCS (5.3%).
- Mayor rezago: Guanajuato, Chiapas y Oaxaca (>13%).

Figure A3. Proportion of Individual's Educational Lag by State in Mexico

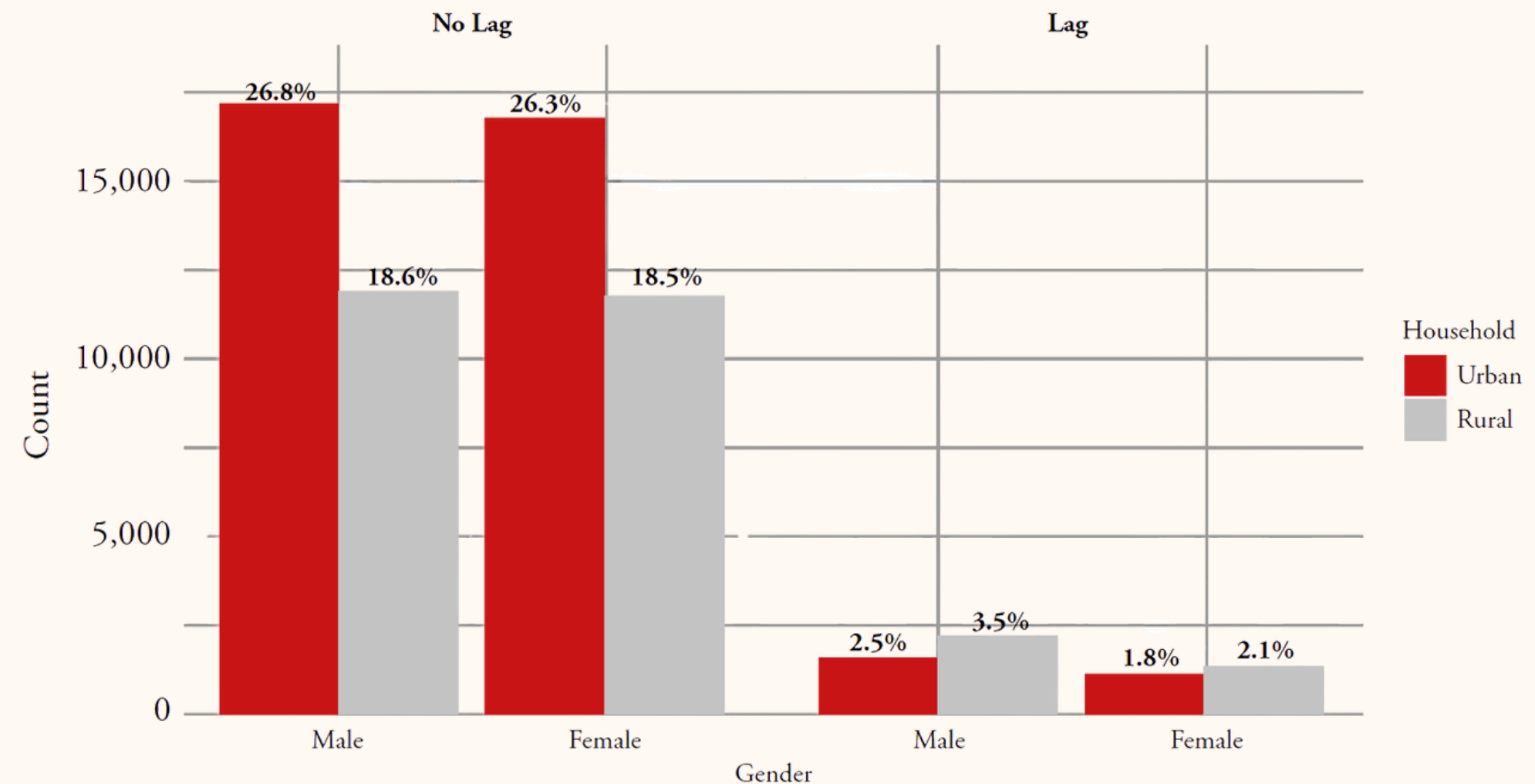


FACTORES ESTRUCTURALES EN EL CONTEXTO MEXICANO (1/2)

Género y hogar

- Las mujeres que participan en el mercado laboral enfrentan **desafíos adicionales** para equilibrar el cuidado y el sustento del hogar (Goldin, 2006).
- El “*child penalty*” implica una **caída en los ingresos laborales** tras el nacimiento del primer hijo, y es más severo en madres solteras (Kleven et al., 2024).
- Las madres solteras enfrentan **condiciones laborales más precarias**, menor acceso a empleo estable y mayor estrés económico (Esquivel, 2024; Khatun & Saadat, 2020).
- La menor **escolaridad materna agrava** la exclusión laboral, y reducen más su participación laboral conforme aumenta el número de hijos (INEGI, 2024).
- La presencia de **hermanos mayores puede reducir la carga** de cuidado y mitigar el rezago; su ausencia lo intensifica (Jaume & Willén, 2021).

Figure 1. Individual's Educational Lag by Gender and Rural/Urban Household



Elaboración propia con datos de la ENIGH, 2022.

Hombres y mujeres en contextos rurales exhiben mayores niveles de rezago educativo que quienes viven en zonas urbanas.

FACTORES ESTRUCTURALES EN EL CONTEXTO MEXICANO (2/2)

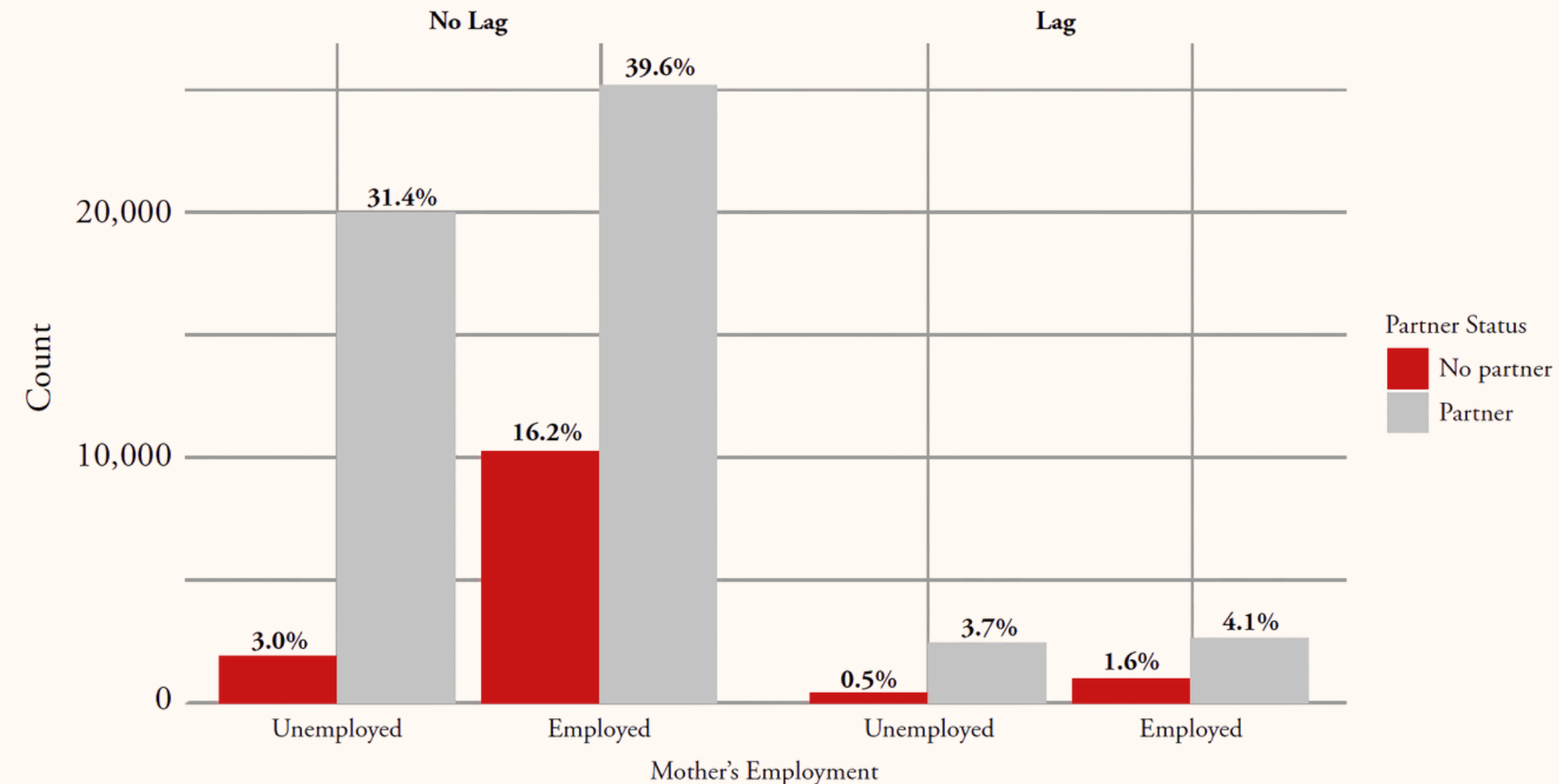
Desigualdad y pobreza

- Existe una conexión entre la **pobreza y el rezago** educativo en México, donde la pobreza perpetúa la desigualdad y limita la movilidad social (Esquivel, 2024).
- Menores recursos en hogares monoparentales detonan **desventajas educativas** (McLanahan y Sandefur, 1994).

Servicios públicos limitados

- La falta de servicios de **cuidado infantil** accesibles y asequibles agrava la carga (López-Estrada, 2017).
- Deficiencias en **seguridad social y vivienda** digna son una amenaza para la educación (INEGI, 2024).

Figure 2. Individual's Educational Lag by Mother's Employment and Partner Status



Elaboración propia con datos de la ENIGH, 2022.

El rezago educativo es más frecuente en hogares donde la madre está empleada y no tiene pareja.

DETERMINANTES DEL REZAGO EDUCATIVO EN MÉXICO

Evidencia de países en desarrollo (Huisman & Smits, 2015)

- Mayor probabilidad de rezago educativo:
 - Tener hermanos (menores y hombres).
 - Hogar monoparental.
 - No tener abuelos.
 - Vivir en zona rural.
 - No tener apego cultural.
 - Transición primaria → secundaria es punto de ruptura.



MÉXICO

Preguntas de investigación

→ **¿CUÁLES SON LOS PRINCIPALES DETERMINANTES DEL REZAGO EDUCATIVO EN MÉXICO?**

→ **¿CÓMO INFLUYEN ESTOS DETERMINANTES EN LA PROBABILIDAD DE REZAGO EDUCATIVO ENTRE LOS HIJOS E HIJAS DE HOGARES MONOPARENTALES ENCABEZADOS POR MUJERES QUE PARTICIPAN EN EL MERCADO LABORAL?**

DETERMINANTES DEL REZAGO EDUCATIVO EN MÉXICO

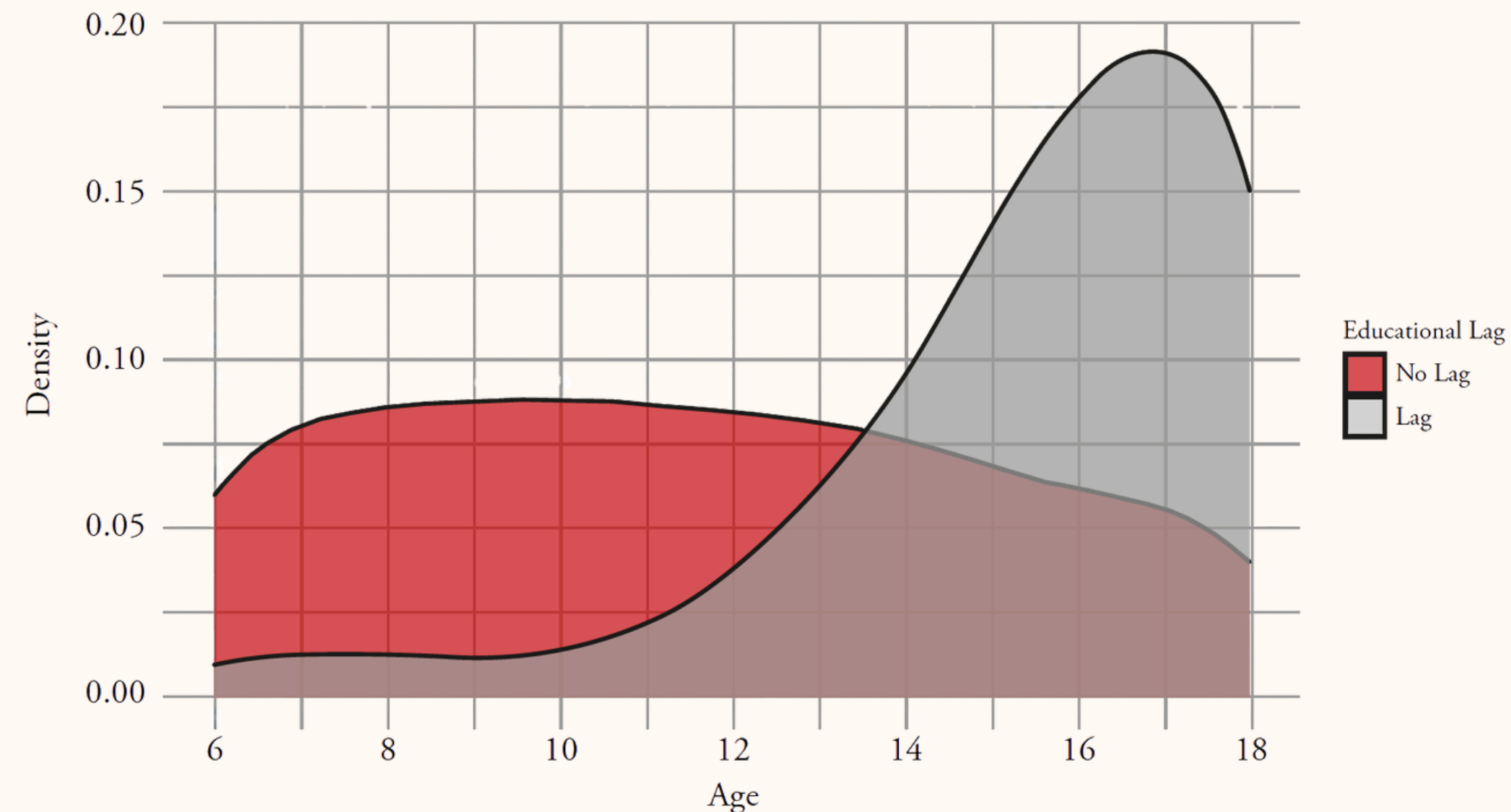
Enfoque

- Niños y jóvenes en hogares monoparentales.
- Madres son jefas de hogar y parte de la fuerza laboral.

Estadística descriptiva

- ENIGH 2022 (Pobreza y Población).
- Muestra: 63,955 observaciones.
 - Jóvenes de 6-18 años que estudian y no trabajan (evita doble causalidad).
- El 9.86% de la muestra presenta rezago educativo.
- El nivel educativo promedio de las madres cuyos hijos presentan rezago es de educación primaria completa; de quienes no presentan rezago, es de secundaria completa.

Figure A1. Individual's Age Distribution by Educational Lag



Elaboración propia con datos de la ENIGH, 2022.

La prevalencia de rezago educativo incrementa con la edad, es más frecuente para jóvenes de 14-18 años.

MARCO METODOLÓGICO

Variable dependiente

- 0: sí tiene el nivel de educación para la edad (no tiene carencia educativa)
- 1: no tiene el nivel de educación para la edad (sí tiene carencia educativa)

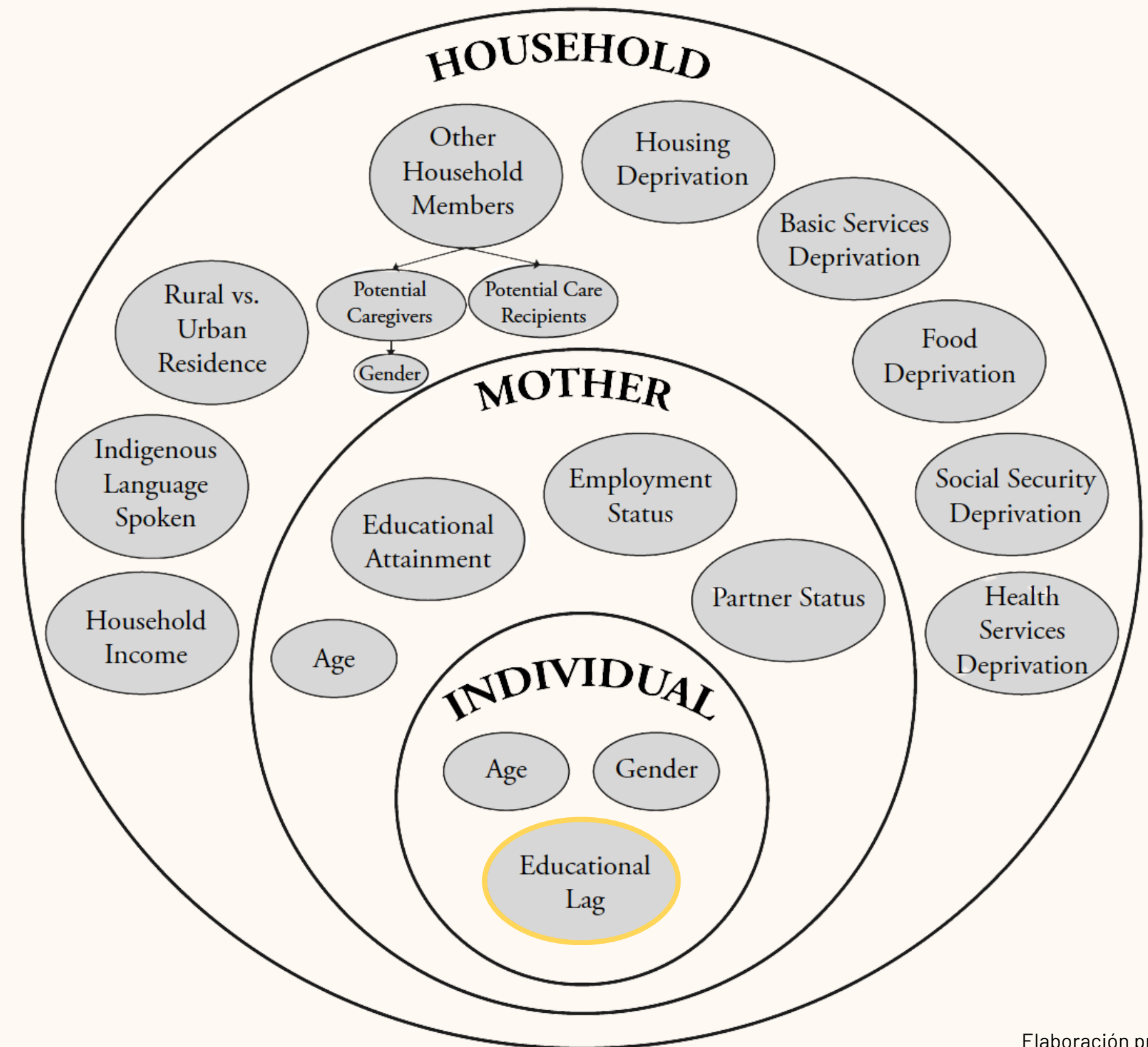
Variables independientes

- Individuales (edad, género).
- Maternas (edad, escolaridad, empleo, pareja).
- Hogar (carencias, ingresos, ruralidad, lengua indígena, otros potenciales cuidadores del hogar).

Marco multidimensional basado en:

Huisman & Smits (2015)
Webbink et al. (2011)
Spierings et al. (2010)

Figure 3. Framework of Variables Grouped by Three Dimensions: Individual, Mother and Household Characteristics



MODELOS

1

PROBIT

- Modela la probabilidad de que un individuo experimente rezago educativo para cada factor explicativo (X_i, M_i, H_i) con base en una función de distribución acumulada normal estándar.

$$Prob\left(education_lag_{X_i} \mid X_i, M_i, H_i\right) = \Phi\left(\beta_0 + \beta_1 X_i + \beta_2 M_i + \beta_3 H_i\right) \quad (1)$$

2

LASSO

- Técnica de regularización que selecciona las variables más relevantes penalizando el tamaño absoluto de los coeficientes de la regresión.

$$\min_{\{\beta_0, \beta_j\}} \left(\sum_{i=1}^n \left(education_lag_{X_i} - \beta_0 - \sum_{j=1}^p \beta_j [X_i, M_i, H_i] \right)^2 + \lambda \sum_{j=1}^p |\beta_j| \right) \quad (2)$$

3

SUPPORT
VECTOR
MACHINES

- Modelo no paramétrico que, mediante funciones *kernel*, identifica fronteras no lineales para clasificar los datos con alta precisión, como complemento a los modelos interpretativos.

$$\min_{\{w, b, \xi_i\}} \frac{1}{2} \|w\|^2 + C \sum_{i=1}^n \xi_i \quad s.t. \quad education_lag_{X_i} \left(w^T [X_i, M_i, H_i] + b \right) \geq 1 - \xi_i, \quad \xi_i > 0 \quad (3)$$

RESULTADOS

Incrementan la probabilidad de rezago educativo

- La edad
- Las carencias
- Un mayor número de personas mayores en el hogar o de cuidadoras mujeres.

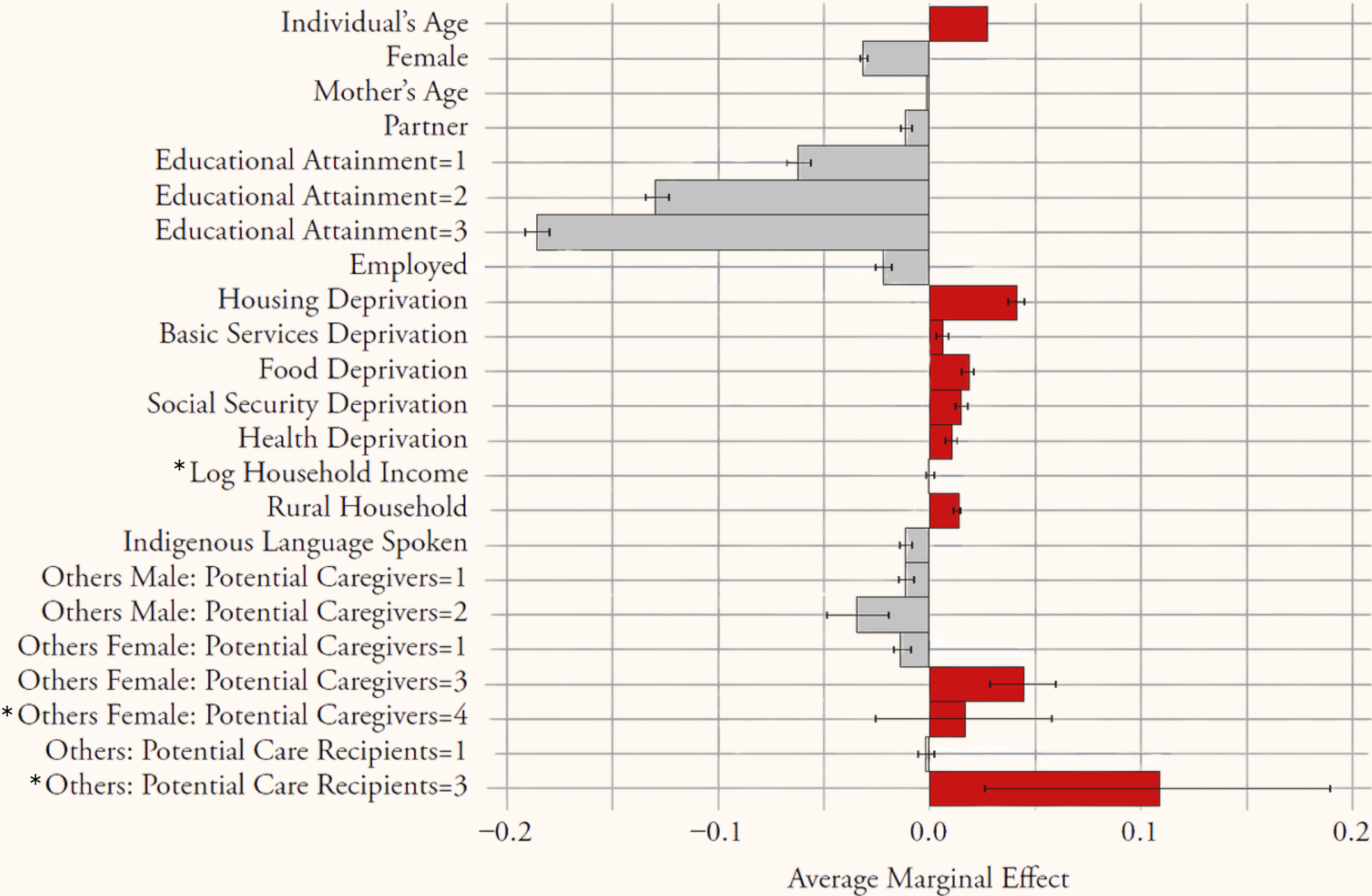
Disminuyen la probabilidad de rezago educativo

- Ser mujer
- Que la madre tenga empleo
- Mayor educación materna
- La presencia de una pareja en el hogar
- El uso de lengua indígena
- Ciertas configuraciones de cuidadores adicionales en el hogar

La mayoría* de los coeficientes son significativos al 99% de confianza.

La dirección y magnitud de los efectos son consistentes con la literatura.

Figure 5. Average Marginal Effects of Probit Model for Individual's Educational Lag



Elaboración propia con datos de la ENIGH, 2022.

HETEROGENEIDAD POR SUBGRUPOS

Edad y educación materna

- De 12 a 15 años ↑ probabilidad de rezago en 19 pp.
- Prob. de rezago ↓ hasta 44 pp para los hijos cuyas madres completaron educación media superior o más.

Condiciones del hogar

- Vivienda inadecuada ↑ prob.de rezago 10–13 pp.
- Carencia alimentaria y de seguridad social ↑ prob. rezago 3–6 pp.
- Pasar de línea de pobreza extrema urbana a Q3 disminuye la probabilidad de rezago sólo en 2–3 pp.

Contexto rural e indígena

- Vivir en zona rural ↑ prob.de rezago 3–4 pp.
- Lengua indígena ↓ prob. rezago 4 pp (contrario a la literatura, posible efecto de cohesión comunitaria).

Cuidadores en el hogar

- 2 cuidadores hombres ↓ prob. rezago hasta 6 pp.
- 4 cuidadoras mujeres ↑ prob. rezago 2–3 pp (por sobrecarga de recursos, rol de hogares extendidos).

Table 3. Probit Marginal Effects: Selected Variables for the Average Male Individual

Subgroup	Individual's Age		Mother's Educational Attainment		Housing Deprivation		Food Deprivation		Social Security Deprivation	
	12	15	0	3	0	1	0	1	0	1
(1) No Partner and Employed	0.12	0.29	0.54	0.11	0.30	0.41	0.31	0.36	0.29	0.34
(2) No Partner and Unemployed	0.13	0.32	0.57	0.13	0.33	0.45	0.34	0.39	0.32	0.37
(3) Partner and Employed	0.10	0.26	0.50	0.10	0.27	0.38	0.28	0.32	0.26	0.31
(4) Partner and Unemployed	0.12	0.29	0.54	0.11	0.30	0.41	0.31	0.36	0.29	0.34

Subgroup	Household Income		Rural Household		Indigenous Language Spoken		Potential Caregivers: Male		Potential Caregivers: Female	
	EIPL	Q3	0	1	0	1	0	2	0	4
(1) No Partner and Employed	0.33	0.31	0.30	0.34	0.33	0.29	0.33	0.27	0.32	0.34
(2) No Partner and Unemployed	0.36	0.34	0.33	0.38	0.36	0.32	0.36	0.30	0.36	0.37
(3) Partner and Employed	0.30	0.28	0.27	0.31	0.30	0.26	0.30	0.24	0.29	0.31
(4) Partner and Unemployed	0.33	0.31	0.30	0.34	0.33	0.29	0.33	0.27	0.32	0.34

Note: These probit model regression results are the average marginal effects of selected characteristics (based on the variables that show the largest effects and higher contrast between the dummies), for the estimation of the individual's educational lag for the average male.

Elaboración propia con datos de la ENIGH, 2022.

Madre sin pareja y sin empleo
Subgrupo (2) presenta la probabilidad más alta de rezago educativo.

Brecha de género persistente
Hombres presentan 20–40% más probabilidad de rezago en todos los subgrupos.

CONCLUSIONES

Limitaciones

- Estructura de corte transversal, no tiene seguimiento longitudinal.
- Posibles sesgos de selección por factores no observables (como el cuidado infantil informal, el apoyo comunitario, etc.).

Hallazgos clave

- No contar con una **vivienda digna** puede incrementar la probabilidad de tener rezago educativo **hasta en un 36%**.
- Ser **hombre** aumenta la probabilidad de tener rezago educativo **entre un 20% y un 40%** en comparación con las mujeres.
 - Los hombres tienen un mayor costo de oportunidad de continuar estudiando que las mujeres.
- La probabilidad de rezago educativo se **reduce hasta en un 77%** entre los hijos de madres que cuentan con al menos **educación media superior**.
 - Existe una representatividad del nivel educativo de la madre ("*role model*") en la permanencia educativa.
- Ser **hablante de lengua indígena** puede reducir la probabilidad de tener rezago educativo **hasta en un 11%**.
 - Dentro de zonas rurales, ser hablante de lengua indígena puede fungir como *proxy* de un mayor apego cultural.

Intervenciones

- Promover la cultura puede resultar en una menor probabilidad de tener rezago educativo.
- Fomentar la permanencia escolar en transiciones críticas, y la flexibilidad laboral y educativa para madres.
- Focalización de apoyos y enfrentar carencias clave: acceso a vivienda digna y seguridad social.



Determinants of Educational Lag in Mexico: A Comprehensive Analysis Using Econometrics and Machine Learning

*Determinantes del rezago educativo en
México: un análisis integral mediante
econometría y aprendizaje automático*

Andrea Fastlicht Perelman
andrea.fastlicht@banxico.org.mx
Banco de México
ORCID: 0000-0001-9705-5080



Fastlicht Perelman, A. (2025).
Determinantes del rezago educativo
en México: Un análisis integral
mediante econometría y aprendizaje
automático. Sobre México Temas De
Economía, 1(12), 146–181.

<https://doi.org/10.48102/rsm.v1i12.172>

¡MUCHAS GRACIAS!

Andrea Fastlicht Perelman
andrea.fastlicht@banxico.org.mx
Banco de México

ANEXO 1

Table 1. Descriptive Statistics: Average by Individual's Gender and Educational Lag

Variable	Female		Male		Summary	
	Lag	No Lag	Lag	No Lag	Min	Max
Observations	2 506	28 612	3 801	29 036		
1. Individual						
Age	15.2	11.6	15.5	11.4	6	18
2. Mother						
Age	41.3	37.8	41.4	37.6	17	75
Partner	0.79	0.79	0.79	0.79	0	1
Educational Attainment	1.22	2.04	1.25	2.06	0	3
Employed	0.57	0.62	0.59	0.62	0	1
3. Household						
Housing Deprivation	0.24	0.12	0.22	0.12	0	1
Basic Services Deprivation	0.36	0.21	0.36	0.21	0	1
Food Deprivation	0.31	0.2	0.3	0.2	0	1
Social Security Deprivation	0.69	0.53	0.7	0.53	0	
Health Services Deprivation	0.48	0.35	0.48	0.35	0	1
Household Income	3 738	4 812	4 131	4 887	0	350 371
Rural Household	0.55	0.41	0.58	0.41	0	1
Indigenous Language Spoken	0.16	0.08	0.15	0.08	0	1
Others Male: Potential Caregivers	0.06	0.07	0.06	0.07	0	3
Others Female: Potential Caregivers	0.54	0.51	0.54	0.5	0	4
Others: Potential Care Recipients	0.12	0.12	0.11	0.12	0	3

Table 2. LASSO and Probit Model: Results for Individual's Educational Lag

Variable	LASSO		Probit	
	Coefficient	Standard Error	Average Marginal Effects	Standard Error
1. Individual				
Age	0.411 ***	0.001	0.027 ***	0.000
Female	-0.459 ***	0.004	-0.031 ***	0.002
2. Mother				
Age	-0.015 ***	0.002	-0.001 ***	0.000
Partner	-0.145 ***	0.002	-0.011 ***	0.003
Educational Attainment = 1	-0.484 ***	0.009	-0.062 ***	0.006
Educational Attainment = 2	-1.249 ***	0.005	-0.129 ***	0.006
Educational Attainment = 3	-2.626 ***	0.004	-0.186 ***	0.006
Employed	-0.246 ***	0.002	-0.022 ***	0.004
3. Household				
Housing Deprivation	0.549 ***	0.006	0.041 ***	0.004
Basic Services Deprivation	0.099 ***	0.005	0.006 *	0.003
Food Deprivation	0.218 ***	0.005	0.018 ***	0.003
Social Security Deprivation	0.252 ***	0.005	0.015 ***	0.003
Health Deprivation	0.116 ***	0.005	0.010 ***	0.003
Log Household Income	-0.014 ***	0.004	0.000	0.002
Rural Household	0.207 ***	0.004	0.013 ***	0.002
Indigenous Language Spoken	-0.143 ***	0.005	-0.011 **	0.003
Others Male: Potential Caregivers = 1	-0.143 ***	0.006	-0.011 *	0.004
Others Male: Potential Caregivers = 2	-0.420 ***	0.035	-0.034 *	0.015
Others Female: Potential Caregivers = 1	-0.132 ***	0.006	-0.013 ***	0.004
Others Female: Potential Caregivers = 3	0.488 ***	0.036	0.044 **	0.016
Others Female: Potential Caregivers = 4	0.200 ***	0.062	0.016	0.042
Others: Potential Care Recipients = 1	-0.030 ***	0.006	-0.002	0.004
Others: Potential Care Recipients = 3	0.556 ***	0.102	0.108	0.082

Source: Author's own elaboration.
Note: Significance levels are the following: * p < 0.10, ** p < 0.05, *** p < 0.01.



ANEXO 2

Figure A2. Household Income Distribution by Individual's Educational Lag

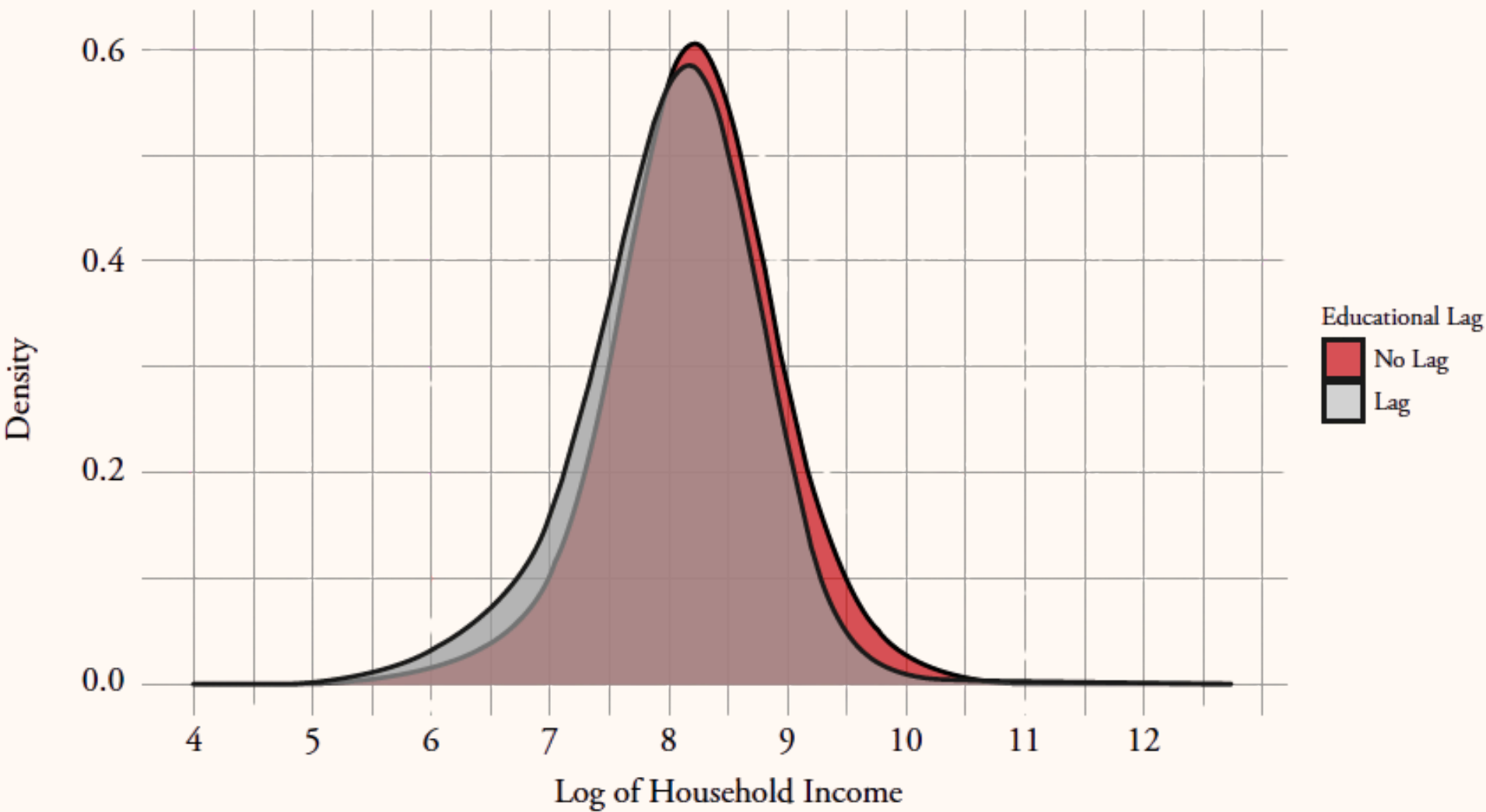
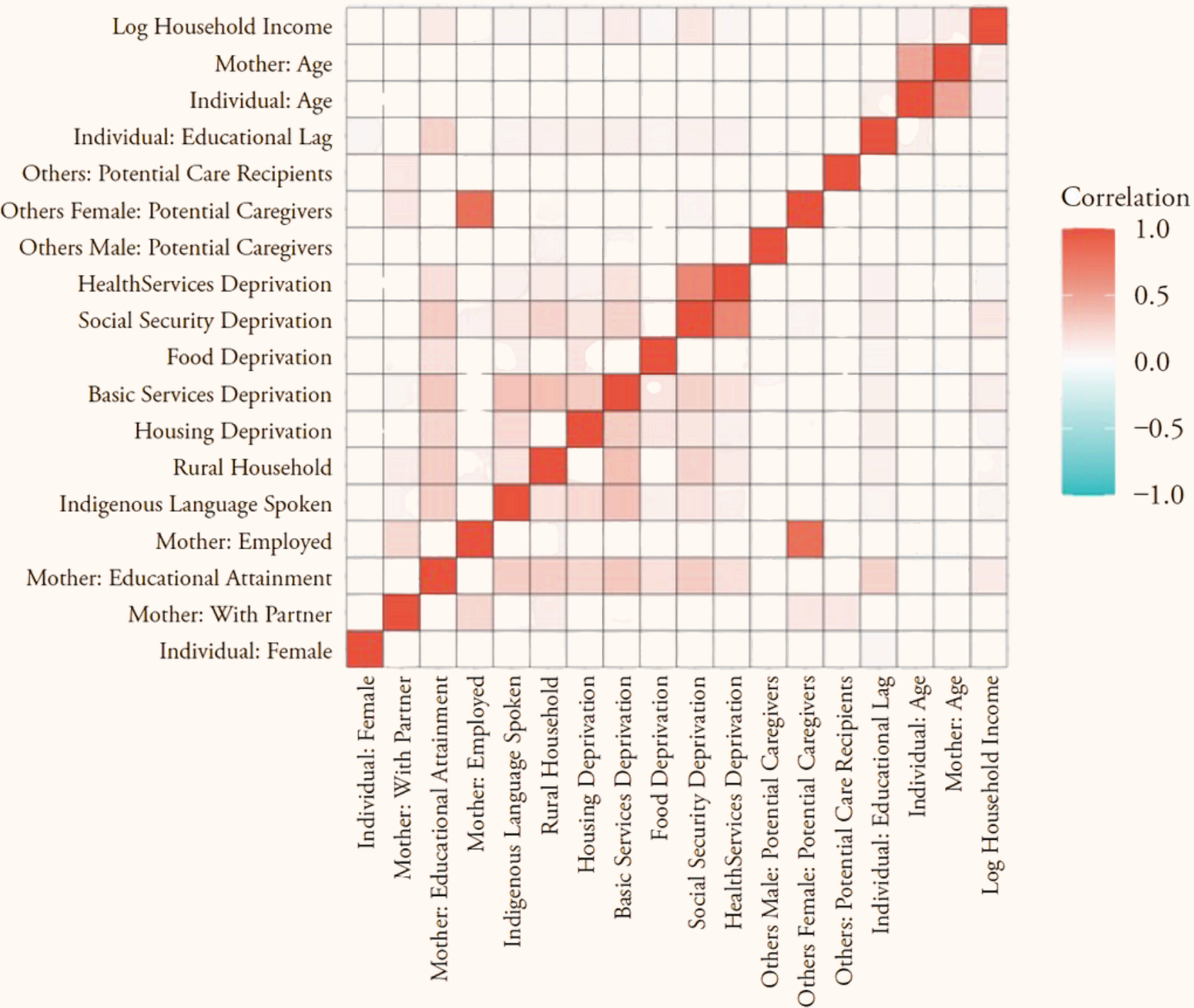


Figure A6. Correlation Heatmap Between Variables



Elaboración propia con datos de la ENIGH, 2022.

Elaboración propia con datos de la ENIGH, 2022.



ANEXO 3

Figure A7. LASSO Cross-Validation Plot for Optimal λ Selection

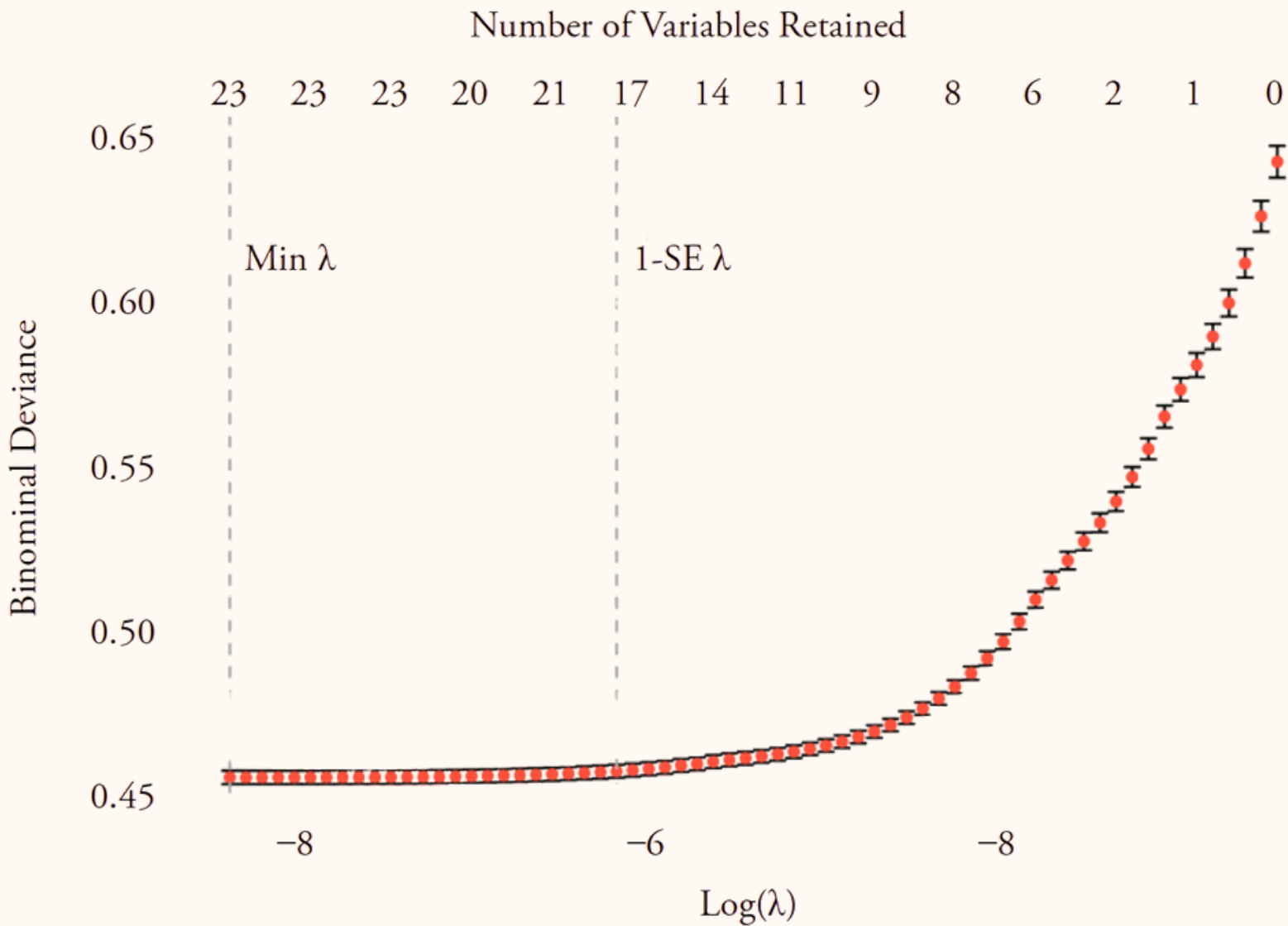
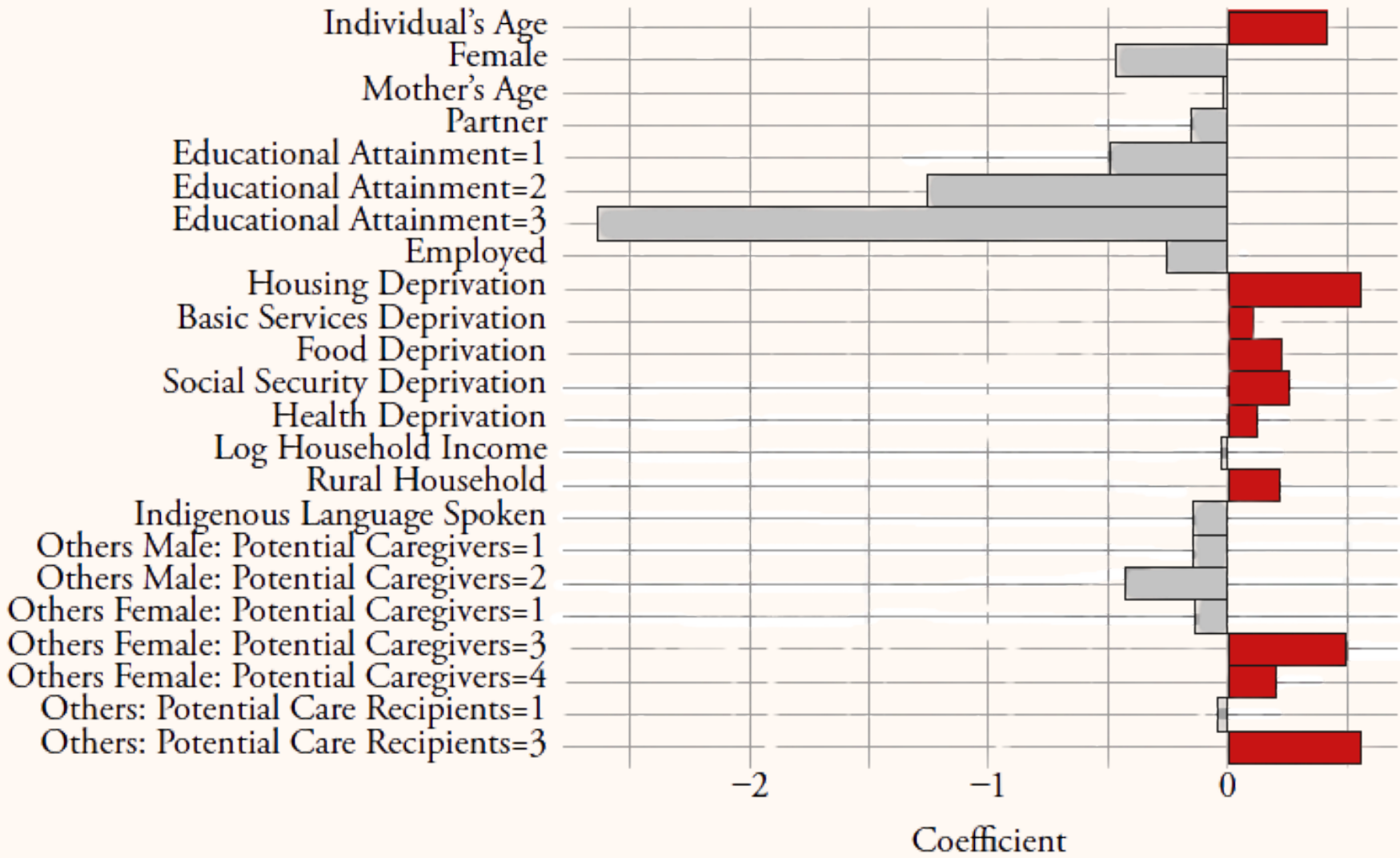


Figure 4. LASSO Regression Coefficients for Individual's Educational Lag



Source: Author's own elaboration.
Note: The minimum λ shows that the number of variables relevant for the model is 23.



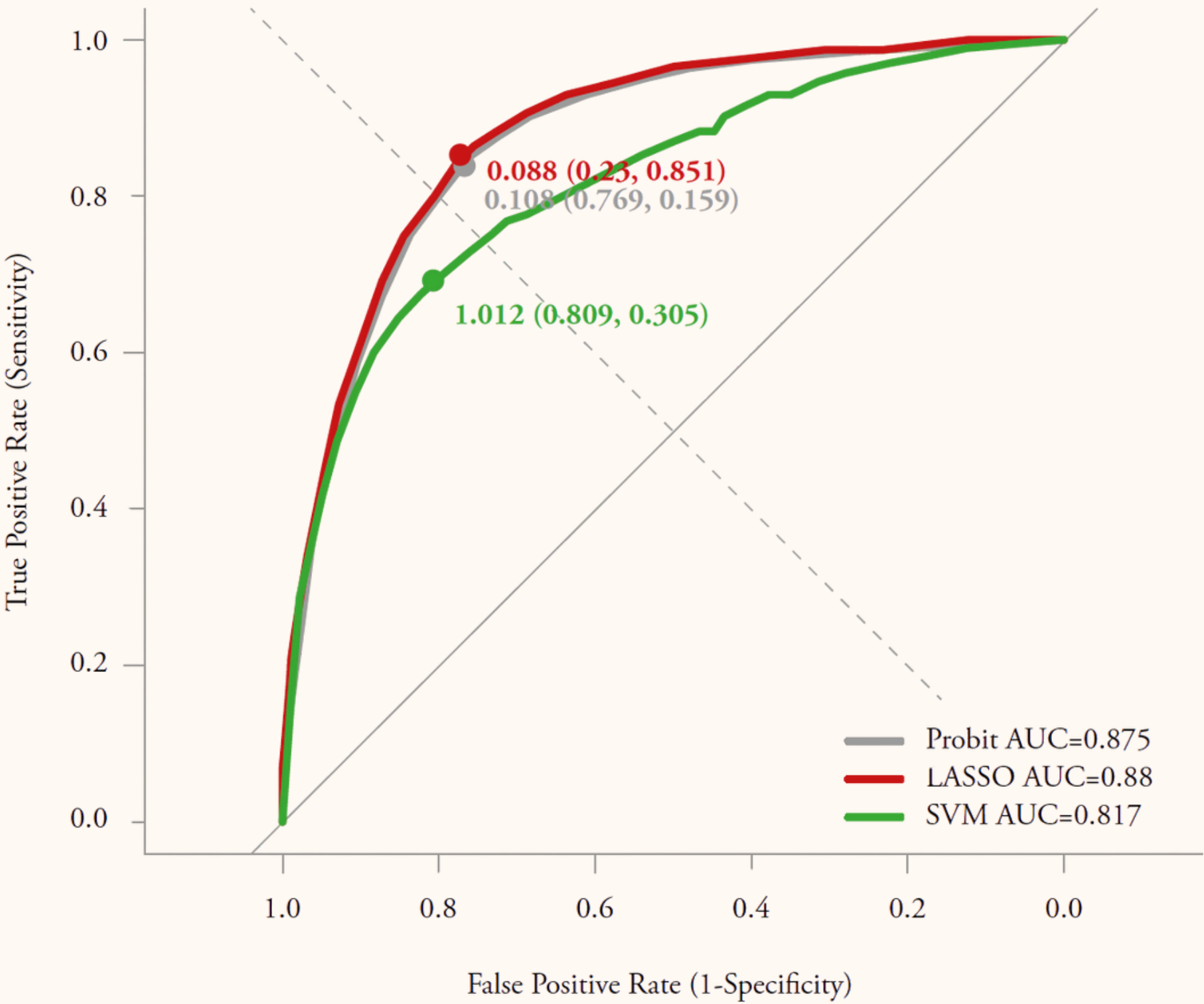
ANEXO 4

Table 4. Probit vs LASSO vs SVM: Performance and Threshold Metrics

Metric	Model		
	Probit	LASSO	SVM
Accuracy	0.790	0.786	0.905
Precision	0.297	0.293	0.673
Sensitivity (True Positive Rate)	0.833	0.838	0.063
Specificity	0.786	0.780	0.997
1-Specificity (False Positive Rate)	0.214	0.220	0.003
Area Under the Curve	0.875	0.880	0.817
Optimal Threshold	0.108	0.088	1.012

Elaboración propia con datos de la ENIGH, 2022.

Figure 7. ROC Curves: Optimal Thresholds for Probit, LASSO and SVM Models



Elaboración propia con datos de la ENIGH, 2022.

