



Brecha salarial de género en México: descomposición con selección muestral y heterogeneidad por formalidad (2018–2025)

Gender Wage Gap in Mexico: Oaxaca–Blinder Decomposition with Sample Selection Correction and Heterogeneity by Employment Formality (2018–2025)

José Alberto Pérez Gómez*[†] y Steve Aguilar Hernández**

Información del artículo	Resumen
Recibido: 06 mayo 2025	Objetivo: Analizar la magnitud, composición y heterogeneidad de la brecha salarial de género en México durante el período 2018–2025, excluyendo 2020 y 2021, con el propósito de evaluar su estabilidad temporal y su heterogeneidad entre los sectores formal e informal del mercado laboral. Metodología: Se aplicó la descomposición Oaxaca–Blinder con corrección de Heckman e identificación paramétrica mediante restricciones de exclusión de fertilidad sobre microdatos deduplicados de la Encuesta Nacional de Ocupación y Empleo (ENOE). Los errores estándar se estimaron mediante el método delta con conglomerados a nivel de unidad primaria de muestreo. Resultados: Los resultados documentaron una brecha salarial condicional estable en torno al 30% durante el período, con un componente explicado negativo que indica una composición observable más favorable para las mujeres. La disparidad se concentró en el sector informal, donde la brecha fue aproximadamente tres veces mayor que en el sector formal. Asimismo, el sesgo de selección fue estadísticamente detectable, aunque económicamente modesto. Implicaciones: Los hallazgos sugieren que la formalización del empleo constituye un mecanismo plausible
Aceptado: 19 junio 2026	
Clasificación JEL: C31; C34; J16; J31; J71.	
Palabras clave: brecha salarial de género; descomposición Oaxaca–Blinder; corrección de Heckman.	

* Facultad de Ciencias Políticas y Sociales de la Benemérita Universidad Autónoma de Puebla, jose.perezgomez@correo.buap.mx, <https://orcid.org/0000-0001-7156-5592>.

[†] Autor de correspondencia.

** Facultad de Economía de la Benemérita Universidad Autónoma de Puebla, steve.aguilar.hdz@gmail.com, <https://orcid.org/0009-0003-4742-9584>.

ISSN Electrónico: 2448-8402 | ISSN Impreso: 1870-221X | ©2026 Los autores



para reducir la brecha salarial agregada, al limitar el espacio en el que operan penalizaciones salariales no atribuibles a dotaciones observables. **Originalidad:** El estudio aporta evidencia sobre la estabilidad temporal y la heterogeneidad de la brecha salarial de género en México mediante la combinación de la descomposición Oaxaca-Blinder con corrección de Heckman, incorporando una comparación entre los sectores formal e informal del mercado laboral.

Article information	Abstract
Received: 06 May 2025	Objective: To analyze the magnitude, composition, and heterogeneity of the gender wage gap in Mexico during the period 2018–2025, excluding 2020 and 2021, with the aim of evaluating its temporal stability and differences across the formal and informal labor market sectors. Methodology: An Oaxaca–Blinder decomposition with Heckman selection correction and parametric identification based on fertility exclusion restrictions was applied to deduplicated microdata from the National Survey of Occupation and Employment (ENOE). Standard errors were estimated using the delta method with clustering at the primary sampling unit level. Findings: The results documented a conditional wage gap that remained stable at around 30% throughout the period, with a negative explained component indicating that women exhibited a more favorable observable composition on average. The disparity was concentrated in the informal sector, where the gap was approximately three times larger than in the formal sector. Selection bias was statistically detectable but economically modest. Implications: The findings suggest that employment formalization constitutes a plausible mechanism for reducing the aggregate gender wage gap by limiting the scope for wage penalties not attributable to observable endowments. Originality: The study provides evidence on the temporal stability and heterogeneity of the gender wage gap in Mexico by combining the Oaxaca–Blinder decomposition with Heckman selection correction and explicitly comparing formal and informal labor market sectors.
Accepted: 19 June 2026	
JEL Classification: C31; C34; J16; J31; J71.	
Keywords: gender wage gap; Oaxaca–Blinder decomposition; Heckman correction.	

Introducción

La brecha salarial entre hombres y mujeres representa una forma de desigualdad económica en las sociedades contemporáneas, a pesar de los avances educativos, las disparidades observadas exceden las diferencias en capital humano. Los rasgos estructurales del mercado laboral mexicano están condicionados por la magnitud del empleo informal, que representa la mitad de la fuerza laboral ocupada, la heterogeneidad regional y el reciente choque pandémico de COVID-19, que perturbó las trayectorias laborales de manera diferenciada entre hombres y mujeres. El presente estudio aborda la pregunta de cuán amplia, cómo se compone

y dónde se concentra la disparidad salarial de género en México durante el período 2018-2025.

La literatura aplicada sobre la brecha salarial en México ha producido evidencia con énfasis y alcances heterogéneos. Trabajos como los de Cuéllar y Moreno (2022) y Rodríguez (2018) documentan la magnitud agregada de la disparidad para distintos períodos pre pandémicos, situándola entre 25 y 40% según especificación y muestra. Castro Lugo *et al.* (2021) examinan la convergencia regional de la brecha y sugieren que su reducción nacional ha estado impulsada por dinámicas dentro del sector informal. Rosales *et al.* (2023) muestran que la precariedad laboral amplifica las disparidades de género en sectores específicos. Romero Martínez y Prudencio Vázquez (2022) abordan la dimensión sectorial con base en datos trimestrales recientes. Estas contribuciones comparten ciertas limitaciones que el presente trabajo busca subsanar, a partir de tres preguntas concretas. Primero, ¿cuál es la magnitud y composición de la brecha salarial de género en México durante 2018-2025, y cómo se distribuyen sus componentes explicado y no explicado? Segundo, ¿la brecha se modificó significativamente a lo largo del período, en particular ante la disrupción pandémica y la recuperación posterior? Tercero, ¿de qué manera la coexistencia de los sectores formal e informal en el mercado laboral mexicano media la disparidad observada?

La estrategia metodológica combina cuatro elementos. En primer lugar, se utiliza la descomposición de Oaxaca-Blinder (Oaxaca, 1973; Blinder, 1973) con un vector de coeficientes de referencia estimado mediante un modelo agrupado (Oaxaca y Ransom, 1994b), implementada a través del algoritmo de Jann (2008). En segundo lugar, se corrige el sesgo de selección muestral mediante el modelo de Heckman (1979) con restricciones de exclusión en función de la tenencia y número de hijos, para identificar paramétricamente la ecuación salarial. En tercer lugar, los errores estándar se estiman mediante el método delta con corrección por conglomerados a nivel de unidad primaria de muestreo, evitando la inflación artificial de la significancia que se produce al utilizar los factores de expansión como ponderaciones (Cameron y Miller, 2015). En cuarto lugar, la base de datos se construye con deduplicación jerárquica del diseño rotativo de la ENOE para los seis años analizados (2018, 2019, 2022, 2023, 2024 y 2025), excluyendo 2020 y 2021 por inconsistencias metodológicas asociadas al levantamiento durante la pandemia.

Los resultados del estudio revelan una realidad contrastante: aunque las mujeres en México superan a los hombres en preparación académica, no

logran eliminar una brecha impulsada por factores estructurales y sesgos que penalizan a la mujer, ante situaciones como la unión conyugal o el trabajo en el sector servicios. El elemento fundamental de la brecha se encuentra en la informalidad, donde la falta de regulación permite que las valoraciones asimétricas operen con mayor fuerza. La desigualdad no es un problema de capacidad, sino de una estructura laboral que requiere mayor formalización para garantizar justicia salarial.

El trabajo se organiza de la siguiente manera: la sección 1 presenta una revisión de la literatura relevante; la sección 2 describe la fuente de datos, la construcción de la muestra y las variables del modelo; la sección 3 detalla la estrategia econométrica; la sección 4 presenta y discute los resultados; por último, se presenta la síntesis de los hallazgos, y se señalan las limitaciones interpretativas del análisis.

1. Marco teórico de las brechas salariales

La teoría del capital humano constituye el marco analítico estándar para el estudio de los determinantes salariales. Mincer (1974) formaliza la ecuación log-lineal del ingreso en función de la escolaridad y la experiencia cuadrática, conocida como ecuación minceriana, esta es la base de la mayoría de las estimaciones empíricas posteriores de retornos al capital humano; por su parte, Polachek (1981) extiende el enfoque al ámbito de la decisión ocupacional y demuestra que las mujeres anticipan interrupciones laborales asociadas a la maternidad, lo que las lleva a autoseleccionarse hacia ocupaciones con menores retornos salariales.

La existencia de brechas salariales sistemáticas entre grupos demográficamente comparables ha sido analizada desde dos tradiciones teóricas complementarias; por un lado, Becker (1957) propone la teoría de la discriminación basada en preferencias, en la cual los grupos discriminados reciben salarios inferiores a su productividad marginal; por otro lado, Phelps (1972) señala que los empleadores recurren a características observables del grupo de pertenencia como señal probabilística de productividad, lo que genera brechas incluso entre individuos con dotaciones idénticas. En ambas tradiciones encontramos que, a pesar de controlar por características productivas observables, el componente residual de la brecha salarial persiste.

1.1 La descomposición Oaxaca-Blinder y sus extensiones metodológicas

Oaxaca (1973) y Blinder (1973) formalizan, de manera independiente pero simultáneamente, el procedimiento estándar para descomponer brechas salariales entre dos grupos. Demostraron que la diferencia salarial promedio puede separarse en una parte explicada por diferencias en las dotaciones observables y otra parte no explicada atribuible a diferencias en cómo el mercado valora esas dotaciones. Esta separación constituye la herramienta canónica de la economía laboral aplicada al estudio de disparidades entre grupos.

Oaxaca y Ransom (1994b) señalan que la elección del vector de coeficientes que se utiliza como referencia no discriminatoria puede alterar la brecha entre sus componentes; por tal motivo, recomiendan el uso de un modelo agrupado (*pooled*). En este sentido, Jann (2008) desarrolla el comando Oaxaca para Stata, que implementa las variantes principales del método, incorpora distintas opciones de referencia y reporta errores estándar.

Fortin (2008), establece que el componente no explicado no sólo recoge la discriminación, además encontramos la heterogeneidad inobservable entre los grupos, errores de medición en los ingresos auto-reportados y errores de especificación en la forma funcional asumida; la autora recomienda referirse a este componente como diferencial no atribuible a dotaciones observables y reservar el término discriminación para situaciones donde la evidencia adicional lo respalde.

En la literatura reciente encontramos a Mussa (2018), quien adapta el método a modelos *double hurdle* con aplicación al gasto parental en educación. Kröger y Hartmann (2021) extienden el método a datos panel mediante una formulación dinámica que aprovecha la dimensión longitudinal de las encuestas de hogares. Otsu y Tanaka (2022) proponen una inferencia basada en verosimilitud empírica que permite contrastes robustos sin supuestos paramétricos sobre la distribución de los errores. Oaxaca y Sierminska (2023) establecen el vínculo formal entre la descomposición Oaxaca-Blinder y la descomposición Kitagawa, lo que abre nuevas posibilidades de aplicación en demografía y epidemiología.

Ospino *et al.* (2010) documentan la evolución del método en aplicaciones de economía laboral, salud, educación y desigualdad regional. Rahimi y Hashemi Nazari (2021) ofrecen una representación gráfica detallada de la descomposición que ha facilitado su enseñanza y difusión. En conjunto, la descomposición Oaxaca-Blinder constituye una herramienta consolidada y en activa evolución metodológica, cuya aplicación rigurosa requiere

atender tanto la elección del grupo de referencia como la interpretación cautelosa del componente no explicado.

1.2 La corrección por sesgo de selección de Heckman en el marco Oaxaca-Blinder

La descomposición clásica asume que la muestra observada es aleatoria, sin embargo, la decisión de participar difiere entre hombres y mujeres. Heckman (1979) propone un modelo en dos etapas que corrige este sesgo, primero estima la probabilidad de participación mediante un *probit* y luego incorpora el inverso de la *ratio* de Mills como regresor adicional en la ecuación salarial.

Puhani (2000), mediante simulaciones Monte Carlo demuestra que la identificación paramétrica del modelo requiere al menos una variable que afecte la decisión de participar, pero esté excluida de la ecuación salarial. En este sentido, Mulligan y Rubinstein (2008) utilizan indicadores de fertilidad como restricciones de exclusión en su análisis de la evolución de la brecha en Estados Unidos, y demuestran que la corrección de selección altera las estimaciones del componente no explicado; por su parte, Olivetti y Petrongolo (2008) replican esta estrategia en un análisis comparativo internacional sobre países OCDE y encontraron que en países con alta participación laboral femenina la corrección reduce la brecha estimada, mientras que en países con baja participación la corrección la incrementa. Jung, Choe y Oaxaca (2022) analizaron las condiciones de empleo riesgoso frente al empleo seguro. Encontraron que la elección de empleo seguro contribuye al componente residual. Jung (2014) analizó previamente este enfoque mediante un modelo de selección basado en actitudes hacia el riesgo y reporta que la heterogeneidad de aversión al riesgo entre hombres y mujeres explica una fracción significativa del componente no explicado tradicional.

Cameron y Miller (2015), por su parte, sistematizan el tratamiento de errores estándar robustos en encuestas con diseños muestrales complejos y recomiendan la corrección por conglomerados a nivel de unidad primaria de muestreo, ya que infla artificialmente la significancia estadística. En conjunto, la corrección de Heckman se ha consolidado como componente estándar de las descomposiciones salariales con corrección de selección, siempre que se identifique al menos una restricción de exclusión válida y se reporten errores estándar robustos al diseño muestral.

1.3 Evidencia empírica de la brecha salarial de género en México

En México, la aplicación del método Oaxaca-Blinder al estudio de la brecha salarial de género ha sido abordado desde distintos enfoques con alcances heterogéneos. Por ejemplo, Popli (2013) analiza los diferenciales salariales por género en México durante 1996-2006 mediante un enfoque distribucional que compara la brecha a lo largo de toda la distribución salarial y descompone por sector formal e informal. La autora encontró que la brecha resulta mayor en el sector formal que en el informal. Arceo-Gómez y Campos-Vázquez (2014) examinan la evolución de la brecha con datos de los censos de 1990, 2000 y 2010. Sus hallazgos muestran una reducción de la brecha promedio, un patrón de pisos pegajosos estables y techos de cristal decrecientes, y un fenómeno de selección positiva de las mujeres al mercado laboral.

Rodríguez Pérez (2018) analiza la brecha salarial regional según la exposición a la apertura comercial durante 2005-2015 encontrando heterogeneidad entre regiones del país. Posteriormente, Rodríguez (2018) descompone el diferencial salarial por género entre el sector público y el sector privado formal-informal encontrando evidencia directa sobre el papel de la informalidad en la brecha. En este sentido, Castro Lugo *et al.* (2021) analizan la convergencia regional de la brecha entre estados mexicanos durante 2005-2017. Concluyen que la reducción nacional estuvo impulsada por dinámicas dentro del sector informal. Por su parte, Rosales *et al.* (2023) aplican el análisis al sector turístico mexicano y encontraron que la precariedad laboral de la informalidad en el sector, amplifica las disparidades de género.

Romero Martínez y Prudencio Vázquez (2022) aplican la descomposición estándar a la ENOE de 2013 y 2018, encontrando persistencia de la brecha durante ese período. En este sentido, Cuéllar y Moreno (2022) construyen un análisis temporal más extenso para México (1988 - 2019), homologando las encuestas ENEU, ENE y ENOE, para un modelo Heckman seguido de una extensión Blinder-Oaxaca tipo Mulligan-Rubinstein. Encontraron que la brecha fluctúa alrededor del 29.6% una vez corregido el sesgo de selección, mientras que la omisión de dicha corrección la subestima en aproximadamente 6.7 puntos porcentuales en promedio.

Arceo-Gómez, Campos-Vázquez, Medina y Vélez (2018) implementan un experimento sobre negociación y preferencias económicas por género en México. Encontraron que las diferencias en la disposición a negociar se

traducen en brechas salariales no atribuibles a dotaciones. Canedo (2019), Arceo-Gómez y Torres (2021), y De la Luz Tovar y Pineda Salazar (2023) aplican el método a brechas salariales por autoidentificación indígena en México. Sus hallazgos son similares a los observados por género, pero con componentes no explicados sustanciales y heterogéneos por región.

En el contexto latinoamericano, Salce Díaz (2021) analiza la evolución de la brecha salarial por género en Chile y encuentra una reducción gradual durante las últimas dos décadas. Por otro lado, Linthon-Delgado y Méndez-Heras (2022) reportan un componente no explicado que supera el 70% de la brecha total para Ecuador. Además, Bonavida y Gasparini (2024) argumentan que la composición ocupacional en América Latina explica una fracción minoritaria de las disparidades observadas, mientras que la mayor parte permanece como componente no explicado.

De los estudios anteriores se desprende que la brecha salarial de género en México es persistente, geográficamente heterogénea y dominada en magnitud por su componente no explicado, con la informalidad apareciendo de manera recurrente como factor estructural en su reproducción.

La literatura revisada presenta cinco aspectos en los que el presente estudio busca aportar a la evidencia acumulada sobre la brecha salarial de género en México. En primer lugar, tenemos la ampliación del horizonte temporal 2018-2025 (excluyendo 2020 y 2021). Los años post pandémicos permiten evaluar la estabilidad estructural de la brecha frente al choque económico más significativo de las últimas décadas. En segundo lugar, estimamos descomposiciones Oaxaca-Blinder por separado para los sectores formal e informal con identificación paramétrica válida del modelo Heckman. En tercer lugar, en el diseño muestral implementamos una deduplicación jerárquica explícita basada en los identificadores de hogar, vivienda y persona, con la finalidad de evitar inflar artificialmente los tamaños muestrales. Además, permite reforzar la consistencia de los errores estándar reportados. Finalmente, encontramos una evolución del parámetro λ , que transita de signo negativo en 2018 a signo positivo en 2022-2025. Eso sugiere una transformación post pandémica del proceso de selección laboral femenina; sin embargo, el resultado no alcanza significancia estadística suficiente para sostener una afirmación causal y se reserva como hipótesis para investigación futura.

2. Datos

Los datos fueron tomados de la Encuesta Nacional de Ocupación y Empleo (ENOE), proporcionada por el Instituto Nacional de Estadística y Geografía (INEGI). El período se conformó de seis años: 2018, 2019, 2022, 2023, 2024 y 2025, se excluyeron los años 2020 y 2021 debido a que, en el 2020, INEGI sustituyó temporalmente la ENOE presencial por la Encuesta Telefónica de Ocupación y Empleo (ETOE) ante las restricciones sanitarias de la pandemia. Por su parte, en el 2021 inició la transición operativa hacia la ENOE Nueva Edición, con ajustes en el marco muestral y en los procedimientos de levantamiento, por lo que la cobertura y el diseño podrían generar inconsistencias que comprometerían su comparabilidad.

En la construcción de la base de datos se consideró el diseño rotativo de la ENOE. En la mecánica de la encuesta cada vivienda permanece cinco trimestres consecutivos en la muestra de la ENOE, al apilar los cuatro trimestres de un año sin filtros adicionales, el mismo individuo aparecerá hasta cuatro veces, inflando artificialmente el tamaño muestral y subestimando los errores estándar al violar el supuesto de independencia entre observaciones. Para evitarlo se aplica una deduplicación jerárquica sobre la combinación (cd_a, ent, con, v_sel, n_hog, h_mud, n_ren), conservando la primera aparición de cada persona en el año. Tras este filtro, los tamaños muestrales son: 181,006 (2018), 117,450 (2019), 192,464 (2022), 213,409 (2023), 233,872 (2024) y 132,950 (2025), criterio coherente con Romero Martínez y Prudencio Vázquez (2022) y Andrés Rosales, Mendoza y Quintana (2023) quienes emplean el primer trimestre de cada año. Por su parte, Rodríguez (2018) utiliza únicamente el tercer trimestre, señalando que esto también evita efectos estacionales.

A partir de estas consideraciones, se construyó una base de datos individual para hombres y mujeres en edad laboral. La variable dependiente es el logaritmo del ingreso mensual por ocupación principal¹, condicional en su percepción positiva, en línea con la formulación minceriana clásica (Mincer, 1974) y la práctica estándar en los estudios de brecha salarial de género en México (Cuéllar y Moreno, 2022; Castro Lugo *et al.*, 2021). Las variables explicativas se clasifican en continuas

¹ Construido a partir de la variable *ingocup* de la ENOE, que reporta el ingreso monetario mensual derivado de la ocupación principal; se descartan observaciones con ingreso cero o ausente.

como la *experiencia potencial*² y su cuadrado, y dicotómicas, en este grupo se consideran dummies por *nivel educativo*³, un indicador de *analfabetismo*⁴, tres indicadores de *estado civil*⁵, dos dummies *sectoriales*⁶, la condición de *formalidad*⁷ y *subocupación*⁸, el ámbito geográfico distinguiendo *zonas rurales*⁹ y residencia en estados fronterizos del *norte*¹⁰, para la ecuación de selección del modelo Heckman, se utilizaron dos indicadores de fertilidad¹¹. La elección responde, por un lado, a la literatura estándar sobre ecuaciones mincerianas de salarios y, por otro, a la disponibilidad de información homogénea en la ENOE durante el período analizado.

La tabla 1 resume las medias y desviaciones estándar de las variables del modelo, diferenciadas por sexo y condición de ocupación. Se observan asimetrías de género que condicionan la oferta de trabajo. Destaca que las mujeres ocupadas promedian 10.67 años de escolaridad frente a los 9.97 de los hombres y muestran una mayor concentración en los niveles educativos superiores (26.4% con educación universitaria y 2.3% con posgrado, en comparación con el 18.7% y 1.4% de los hombres). Eso evidencia que las mujeres con mayor acumulación de capital humano logran sortear las barreras de entrada al mercado formal. Sin embargo, la cualificación no se refleja en paridad de ingresos, ya que, el salario promedio mensual de los hombres es de \$9,115 frente a los \$7,175.07 de

² Calculada de forma minceriana como la edad menos los años de escolaridad menos seis (años previos al inicio de la vida escolar), truncando en cero los valores negativos. Se incluye también su cuadrado dividido entre cien para capturar la concavidad esperada del perfil edad-salario.

³ Seis dummies que toman valor 1 según el máximo nivel educativo alcanzado: *sin estudios* (referencia), *básica* (primaria), *secundaria*, *media superior* (preparatoria), *superior* (universidad) y *posgrado*.

⁴ Toma valor 1 si la persona declara no saber leer ni escribir, y 0 en caso contrario.

⁵ Tres categorías mutuamente excluyentes, *unida* (casada o en unión libre), *separada* (separada, divorciada o viuda) y *soltera* (categoría de referencia).

⁶ Dos dummies que identifican la *rama de actividad*, *secundaria* (industria) y *terciaria* (servicios); la actividad primaria funciona como categoría de referencia.

⁷ Toma valor 1 si la persona ocupada cuenta con acceso a instituciones de seguridad social vinculadas a su empleo, y 0 si no.

⁸ Toma valor 1 si la persona reporta disponibilidad y necesidad de trabajar más horas, y 0 en caso contrario.

⁹ Toma valor 1 si la persona reside en localidad rural (menos de 2,500 habitantes) y 0 si reside en localidad urbana.

¹⁰ Toma valor 1 si reside en alguno de los seis estados fronterizos del norte (Baja California, Coahuila, Chihuahua, Nuevo León, Sonora o Tamaulipas) y 0 en el resto del territorio nacional.

¹¹ La variable *tiene_hijos* toma valor 1 si la mujer reporta al menos un hijo nacido vivo y 0 en caso contrario; *hijos_3mas* toma valor 1 si reporta tres o más. Ambas funcionan como restricciones de exclusión en la ecuación de selección del modelo Heckman.

las mujeres. Además, se aprecia una fuerte concentración de la fuerza laboral femenina en el sector terciario (78.3%).

La población femenina no ocupada, duplica en tamaño a las trabajadoras activas. Además, reportan un promedio de hijos mayor (2.16) que las ocupadas (1.94), con una tasa de incidencia de familias numerosas que alcanza el 36.1% frente al 33.1% del grupo con empleo formal o informal. Por su parte, la variable estado civil muestra que el 64.6% de los hombres ocupados se encuentra casado en contraste con el 50.4% de las mujeres trabajadoras, disparidad que se invierte en la categoría de separadas o divorciadas (17.6% en mujeres ocupadas contra 6.4% en hombres). La evolución anual de estos estadísticos se reporta en el anexo A.

Tabla 1
Estadísticas descriptivas por sexo, ENOE 2018-2025 (datos deduplicados)

	Hombres ocupados	Mujeres ocupadas	Mujeres no ocupadas
Edad	39.049 (14.751)	39.327 (13.756)	41.851 (20.225)
Años de escolaridad	9.978 (5.325)	10.675 (5.368)	9.623 (5.855)
exp_pot	23.161 (16.245)	22.739 (15.627)	26.333 (22.112)
Número de hijos	. (.)	1.940 (1.934)	2.159 (2.619)
salario	9115.866 (8276.207)	7175.073 (6481.055)	. (.)
lsalario	8.855 (0.758)	8.566 (0.847)	. (.)
rural	0.546 (0.498)	0.493 (0.500)	0.470 (0.499)
Estado fronterizo del norte	0.204 (0.403)	0.210 (0.407)	0.179 (0.384)
formal	0.445 (0.497)	0.435 (0.496)	0.482 (0.500)
subocupado	0.085 (0.278)	0.075 (0.264)	0.014 (0.119)
unida	0.646 (0.478)	0.504 (0.500)	0.512 (0.500)
separada	0.064 (0.244)	0.176 (0.380)	0.146 (0.354)
soltera	0.290 (0.454)	0.321 (0.467)	0.342 (0.474)
tiene_hijos	. (.)	0.735 (0.442)	0.673 (0.469)
hijos_3mas	. (.)	0.331 (0.471)	0.361 (0.480)

educ1	0.029 (0.169)	0.028 (0.164)	0.054 (0.226)
educ2	0.209 (0.406)	0.170 (0.376)	0.213 (0.410)
educ3	0.322 (0.467)	0.290 (0.454)	0.291 (0.454)
educ4	0.238 (0.426)	0.225 (0.417)	0.207 (0.405)
educ5	0.187 (0.390)	0.264 (0.441)	0.221 (0.415)
educ6	0.014 (0.116)	0.023 (0.149)	0.013 (0.113)
analfabeta	0.026 (0.160)	0.024 (0.154)	0.050 (0.217)
act_sec	0.336 (0.472)	0.185 (0.388)	0.032 (0.176)
act_ter	0.508 (0.500)	0.783 (0.412)	0.175 (0.380)
Observaciones	319,056	223,398	528,697

Fuente: cálculos del autor con base en la Encuesta Nacional de Ocupación y Empleo (ENOE), INEGI. Período: trimestres apilados de 2018, 2019, 2022, 2023, 2024 y 2025. Se excluyen 2020 (por aplicación de la ETOE durante la pandemia) y 2021 (por la transición a ENOE-N). Las medias están ponderadas por el factor de expansión (fac). Para variables continuas se reporta media con desviación estándar entre paréntesis. Para variables binarias la media se interpreta como la proporción de la categoría. El salario está expresado en pesos corrientes del trimestre de referencia. La muestra restringe a personas de 14 años o más con entrevista válida. 'Ocupados' incluye empleados, subempleados y autoempleados con ingresos positivos. Los puntos (.) indican variables no aplicables al grupo correspondiente.

3. Metodología

Se utilizó la descomposición doble y triple Oaxaca-Blinder¹² para analizar las diferencias salariales entre hombres y mujeres, mediante la corrección por sesgo de selección mediante el modelo de dos etapas de Heckman. Donde las funciones de ingresos fueron estimadas de manera separada, para cada grupo, de igual manera se utilizó un modelo agrupado (*pooled*) como referencia no discriminatoria.

El objetivo central de la descomposición O-B es, por lo tanto, identificar qué proporción de la diferencia en los resultados salariales entre hombres y mujeres se debe a la diferencia en los niveles de las variables explicativas (características individuales o laborales) y qué parte se debe a diferencias en la remuneración entre ambos grupos. Siguiendo la metodología propuesta por Jann (2008) se procede primero a dividir la muestra en dos grupos: Grupo H (hombres) y Grupo M (mujeres). A partir de esta división, se estima una función de ingresos para cada grupo mediante regresiones

¹² (Oaxaca R. , 1973; Blinder, 1973) desarrollaron la metodología para analizar la diferencia entre dos grupos, dividiendo las disparidades en componentes explicables y no explicables.

separadas, donde el salario es la variable dependiente y las características individuales observables las variables independientes. El paso posterior consiste en calcular la descomposición, la cual se desglosa en las dos partes mencionadas anteriormente: el componente explicado por las características observables y el componente residual no explicado

$$\Delta \bar{Y} = \bar{Y}_H - \bar{Y}_M \quad (1)$$

$\bar{Y}_H - \bar{Y}_M$ denotan el salario promedio (en logaritmos) de hombres y mujeres, respectivamente.

Primero se estimó la descomposición Oaxaca-Blinder clásica (Oaxaca, 1973; Blinder, 1973), implementada mediante el algoritmo de Jann (2008) con un coeficiente de referencia no discriminatorio estimado a partir de un modelo agrupado (*pooled*). Divide la diferencia en los resultados medios en una parte explicada por las diferencias entre grupos en las características observables, y otra parte que permanece sin explicar, atribuible a diferencias en los retornos a esas características, de tal forma que reorganizando términos obtenemos la descomposición estándar:

$$\Delta \bar{Y} = (\bar{X}_H - \bar{X}_M)' \hat{\beta}_M + \bar{X}_M' (\hat{\beta}_H - \hat{\beta}_M) \quad (2)$$

La descomposición anterior utiliza los coeficientes del grupo M (mujeres) como referencia, lo cual presupone que la estructura salarial femenina representa el escenario no discriminatorio. Para evitar el problema del grupo de referencia (Oaxaca y Ransom, 1994), el vector de coeficientes de referencia $\hat{\beta}^*$ se estima a partir de un modelo agrupado (*pooled*) que combina ambos grupos en una sola regresión con un indicador de sexo:

$$Y = X\hat{\beta}^* + \epsilon \quad (3)$$

Donde:

X representa la matriz de regresores combinada de ambos grupos

$\hat{\beta}^*$ representa el vector de coeficientes estimado con todas las observaciones.

ϵ representa el error aleatorio

$\hat{\beta}^*$ se obtiene mediante la fórmula general de mínimos cuadrados ordinarios:

$$\hat{\beta}^* = (X'X)^{-1}X'Y \quad (4)$$

Donde:

X es la matriz de las variables explicativas

Y es el vector de la variable dependiente en toda la muestra

El modelo agrupado estima retornos salariales para ambos grupos a la vez, indiscriminadamente. Además, los errores estándar se calculan mediante el método delta a partir de la matriz de varianzas-covarianzas del modelo Heckman de máxima verosimilitud, con corrección por conglomerados a nivel de unidad primaria de muestreo (UPM).

Para la corrección de posibles sesgos por selección muestral, es decir si la decisión de trabajar no es aleatoria, influenciada por factores como educación o hijos, se aplicó el modelo Heckman¹³ dentro de la descomposición Oaxaca-Blinder, para obtener estimaciones salariales más precisas mediante la siguiente ecuación Probit:

$$S^* = Z\gamma + u \quad (5)$$

Donde:

S^* es una variable latente no observada que representa la propensión a participar en el mercado laboral, la participación efectiva S se observa como $S = 1$ si $S^* > 0$ y $S = 0$ en caso contrario.

Z contiene todas las variables explicativas de X^{14} más las restricciones de exclusión (tiene_hijos, hijos_3mas).

γ es un vector de coeficientes asociados a Z.

u representa el término de error normal estándar, se asume $u \sim N(0,1)$ para identificación del modelo Probit.

Una vez que se ha determinado la población con participación laboral, se estima el salario mediante la siguiente regresión lineal:

$$Y = X\beta + \rho\sigma_\epsilon\lambda + \epsilon^{15} \quad (6)$$

Donde:

Y es el logaritmo del salario observado.

¹³ Desarrollado por James Heckman para corregir el sesgo de la muestra, que ocurre cuando la variable dependiente sólo se observa para un subconjunto de la población, generando estimadores sesgados.

¹⁴ El conjunto X incluye experiencia potencial, niveles educativos, indicadores geográficos, estado civil, sector de actividad y formalidad.

¹⁵ La ecuación salarial incorpora la corrección de selección de Heckman a través del *Inverse Mills Ratio* (λ) lo que permite corregir el sesgo derivado de que los salarios sólo se observan para quienes están empleado, el coeficiente ρ mide el impacto de la selección en los salarios mientras que ϵ representa el término de error.

X es un vector de características explicativas (educación, experiencia, etc.).

β son los coeficientes a estimar.

λ es el inverso del Mills *ratio*.

ρ es la correlación entre los errores de las ecuaciones de selección y salario.

σ_ϵ es la desviación estándar del error salarial; el coeficiente que multiplica al inverso de Mills es su producto.

ϵ es el término de error.

La identificación paramétrica del modelo Heckman requiere al menos una variable en el vector Z que no esté contenida en X (restricción de exclusión). Sin esta restricción, el modelo se identifica únicamente por la no linealidad funcional del inverso de Mills, lo cual constituye una identificación débil. Por ello incorporamos como instrumentos de exclusión dos variables que capturan de forma no lineal la carga de cuidado: la primera, *tiene_hijos*¹⁶ y la segunda, *hijos_3mas*¹⁷. La presencia y el número de hijos afectan la decisión femenina de participar en el mercado laboral; sin embargo, no afectan directamente el salario observado condicional en las características de capital humano y posición ocupacional (Mulligan y Rubinstein, 2008; Olivetti y Petrongolo, 2008). En el contexto mexicano, la división sexual del trabajo doméstico sigue siendo marcadamente desigual. La relevancia estadística de las restricciones se verifica mediante un test de Wald conjunto sobre los coeficientes de *tiene_hijos* e *hijos_3mas* en la ecuación de selección. Los resultados rechazan la hipótesis nula de irrelevancia con $p < 0.001$ en cuatro de los seis años analizados, con estadísticos χ^2 entre 32.70 y 143.31. Esto valida la identificación del modelo (resultados detallados en el anexo B). El modelo también asume normalidad bivariada de los errores (u, ϵ), supuesto estándar del Heckman MLE.

En segundo lugar, se estimó la brecha salarial, mediante la descomposición triple de O-B con una función minceriana, (Mincer, 1974). El ingreso de mujeres y hombres depende de variables y los siguientes parámetros:

$$\ln \text{Salario}_H = \beta_H^0 + \sum_{i=1}^n \beta_H^i X_H^i + \epsilon_H \quad (7)$$

¹⁶ Toma valor de 1 si la mujer reporta al menos un hijo.

¹⁷ Toma valor de 1 si reporta tres o más hijos.

$$\ln \text{Salario}_M = \beta_M^0 + \sum_{i=1}^n \beta_M^i X_M^i + \varepsilon_M \quad (8)$$

La diferencia promedio del cambio porcentual de los salarios es:

$$\Delta \ln \text{salario} = \overline{\ln \text{salario}_H} - \overline{\ln \text{salario}_M} \quad (9)$$

Descompuesto en tres componentes:

$$\begin{aligned} \Delta \ln \text{salario} = & (\bar{X}_H - \bar{X}_M)' \hat{\beta}_M + \bar{X}_M' (\hat{\beta}_H - \hat{\beta}_M) \\ & + (\bar{X}_H - \bar{X}_M)' (\hat{\beta}_H - \hat{\beta}_M)^{18} \end{aligned} \quad (10)$$

Donde: $(\bar{X}_H - \bar{X}_M)' \hat{\beta}_M$ mide el efecto en dotación, muestra el cambio de los salarios promedios de un grupo si tuvieran las características del otro; $\bar{X}_M' (\hat{\beta}_H - \hat{\beta}_M)$ mide cuánto cambiarían el salario promedio de un grupo si sus características fueran valoradas como las del otro, $(\bar{X}_H - \bar{X}_M)' (\hat{\beta}_H - \hat{\beta}_M)$ mide el efecto interacción, explica el efecto simultáneo de las diferencias en características y en cómo se valoran esas características entre los grupos.

4. Resultados

Los resultados se presentan en cinco partes que responden de manera secuencial a las preguntas centrales del estudio. En primer lugar, se documenta la magnitud y composición de la brecha salarial de género en México durante 2018-2025 (4.1). En segundo, se identifican las variables que explican cada componente de la descomposición (4.2). Continuamos examinando la heterogeneidad de la brecha entre sectores formal e informal, hallazgo que constituye la principal contribución empírica del trabajo (4.3). En cuarto lugar, se evalúa la estabilidad temporal de la brecha y los resultados del modelo de selección Heckman (4.4). Finalmente se presentan ejercicios de robustez (4.5). Los procedimientos fueron realizados en el software Stata mediante la descomposición de Oaxaca-Blinder (1973) y se siguieron a los modelos implementados por Jann (2008).

¹⁸ La descomposición triple de Blinder-Oaxaca separa la brecha salarial en tres componentes: i) el efecto dotación, que estima cuánto aumentarían los salarios de las mujeres si tuvieran las mismas características que los hombres, ii) el efecto coeficiente, que mide la contribución de la diferencia en retornos a las características, y iii) el efecto interacción, que refleja la combinación de diferencias en dotaciones y coeficientes simultáneamente.

4.1 Brecha global: descomposición Oaxaca-Heckman

Como se anticipó en los estadísticos descriptivos, las mujeres ocupadas en México presentan en promedio mejores credenciales educativas que los hombres ocupados y una alta concentración en el sector terciario. Sin embargo, las mujeres, aún con mejores características observables, reciben salarios sistemáticamente menores. Constituye el punto de partida para interpretar los resultados de la descomposición Oaxaca-Heckman que se reportan en la tabla 2.

Tabla 2
Descomposición Oaxaca-Heckman, México 2018-2025

	2018	2019	2022	2023	2024	2025
overall						
group_1	8.632*** (0.004)	8.695*** (0.004)	8.894*** (0.003)	8.966*** (0.003)	9.099*** (0.003)	9.155*** (0.004)
group_2	8.344*** (0.022)	8.354*** (0.017)	8.582*** (0.013)	8.648*** (0.011)	8.778*** (0.009)	8.816*** (0.015)
difference	0.288*** (0.022)	0.341*** (0.017)	0.313*** (0.013)	0.318*** (0.011)	0.322*** (0.009)	0.340*** (0.015)
	-	-	-	-	-	-
explained	0.066*** (0.003)	0.058*** (0.004)	0.051*** (0.003)	0.048*** (0.003)	0.037*** (0.003)	0.046*** (0.004)
unexplained	0.354*** (0.022)	0.399*** (0.017)	0.364*** (0.013)	0.366*** (0.011)	0.358*** (0.009)	0.385*** (0.015)
Observations	181,006	117,450	192,464	213,409	233,872	132,950

Errores estándar por método delta con corrección por conglomerados (UPM).

Datos deduplicados de ENOE. *** p<0.01, ** p<0.05, * p<0.10.

La brecha salarial promedio entre hombres y mujeres, expresada en logaritmos del ingreso mensual, oscila entre 0.288 en 2018 y 0.340 en 2025, con un valor promedio cercano a 0.317 durante el período analizado. La brecha es altamente significativa en los seis años (estadísticos z entre 13.2 y 35.2, todos con p<0.001), lo que descarta la hipótesis nula de igualdad salarial condicional entre hombres y mujeres. Estos valores son consistentes con estimaciones previas para México durante períodos solapados; Rodríguez (2018) reporta brechas de magnitud similar para 2005-2015, y Castro Lugo *et al.* (2021) encuentran patrones comparables en su análisis regional 2005-2017 en México.

El componente explicado por dotaciones observables toma valores negativos durante todo el período (entre -0.066 en 2018 y -0.046 en 2025) esto nos indica que la diferencia en características observables entre hombres y mujeres ocupados no explica la brecha salarial. Si las

mujeres recibieran los retornos del modelo agrupado por sus dotaciones, ganarían en promedio entre 3.7% y 6.6% más que los hombres, este resultado es consistente con la composición educativa diferenciada que la tabla 1 documenta.

El componente no explicado captura la fracción dominante de la brecha, su magnitud oscila entre 0.354 (2018) y 0.399 (2019), con un valor promedio cercano a 0.372. Este componente recoge no sólo lo que potencialmente correspondería a discriminación salarial, sino también heterogeneidad inobservable entre los grupos, omisión de variables relevantes (calidad educativa, redes de contactos, habilidades blandas, productividad latente), errores de medición y errores de especificación. Por ello, en lo que sigue se utiliza el término "componente no explicado", reservando el término “discriminación” para situaciones donde la evidencia adicional lo respalde. Esta posición, defendida por Oaxaca y Ransom (1994a) y reforzada por Fortin (2008), es metodológicamente necesaria.

De la tabla 2 se desprenden tres observaciones importantes. La primera se refiere a la magnitud de la brecha total, en torno al 30% durante todo el período. Segunda, el signo negativo del componente explicado indica que las mujeres mantienen menores ingresos, a pesar de que poseen mejores atributos observables que los hombres. Tercera, los valores son notablemente estables durante 2018-2025, lo cual sugiere que la desigualdad en el mercado laboral mexicano es un fenómeno estructural y persistente, que no ha mostrado cambios profundos a pesar de los diversos contextos económicos o políticas implementadas durante los seis años observados. Este último punto se examina formalmente en la sección 4.4 mediante un test conjunto de igualdad de la brecha entre años.

4.2 Análisis detallado por componente

La tabla 3 desagrega el componente explicado y el no explicado de la descomposición por las variables del modelo. Esta desagregación permite identificar qué variables están afectando dentro de cada componente y, en consecuencia, dónde se origina realmente la brecha salarial.

Tabla 3
Descomposición Oaxaca-Heckman detallada por componente

	2018	2019	2022	2023	2024	2025
	b	b	b	b	b	b
overall						
group_1	8.632*** (0.004)	8.695*** (0.004)	8.894*** (0.003)	8.966*** (0.003)	9.099*** (0.003)	9.155*** (0.004)

group_2	8.344*** (0.022)	8.354*** (0.017)	8.582*** (0.013)	8.648*** (0.011)	8.778*** (0.009)	8.816*** (0.015)
difference	0.288*** (0.022)	0.341*** (0.017)	0.313*** (0.013)	0.318*** (0.011)	0.322*** (0.009)	0.340*** (0.015)
explained	-0.066*** (0.003)	-0.058*** (0.004)	-0.051*** (0.003)	-0.048*** (0.003)	-0.037*** (0.003)	-0.046*** (0.004)
unexplained	0.354*** (0.022)	0.399*** (0.017)	0.364*** (0.013)	0.366*** (0.011)	0.358*** (0.009)	0.385*** (0.015)
explained						
exp_pot	0.009*** (0.002)	0.003 (0.003)	0.010*** (0.002)	0.005*** (0.002)	0.007*** (0.002)	0.010*** (0.003)
exp_pot2	-0.013*** (0.002)	-0.007** (0.003)	-0.015*** (0.002)	-0.011*** (0.002)	-0.013*** (0.002)	-0.018*** (0.003)
educ2	-0.004*** (0.001)	-0.001* (0.001)	-0.004*** (0.001)	-0.004*** (0.001)	-0.004*** (0.001)	-0.005*** (0.001)
educ3	-0.004*** (0.001)	-0.001 (0.001)	-0.003*** (0.001)	-0.003*** (0.001)	-0.004*** (0.001)	-0.003*** (0.001)
educ4	-0.000 (0.001)	0.001 (0.001)	0.000 (0.001)	-0.000 (0.001)	-0.001 (0.001)	-0.000 (0.001)
educ5	-0.028*** (0.003)	-0.030*** (0.003)	-0.022*** (0.002)	-0.020*** (0.002)	-0.015*** (0.002)	-0.018*** (0.003)
educ6	-0.006*** (0.001)	-0.005*** (0.001)	-0.007*** (0.001)	-0.006*** (0.001)	-0.006*** (0.001)	-0.007*** (0.001)
analfabeta	-0.000 (0.000)	0.000 (0.000)	-0.001** (0.000)	0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)
rural	-0.007*** (0.000)	-0.006*** (0.001)	-0.001*** (0.000)	-0.004*** (0.000)	-0.005*** (0.000)	-0.004*** (0.000)
norte	0.000 (0.000)	-0.001*** (0.000)	0.000 (0.000)	0.000 (0.000)	-0.000 (0.000)	0.000 (0.000)
unida	0.013*** (0.001)	0.015*** (0.001)	0.012*** (0.001)	0.012*** (0.001)	0.012*** (0.001)	0.013*** (0.001)
separada	-0.008*** (0.001)	-0.009*** (0.001)	-0.009*** (0.001)	-0.008*** (0.001)	-0.009*** (0.001)	-0.009*** (0.001)
act_sec	0.063*** (0.002)	0.054*** (0.003)	0.058*** (0.002)	0.052*** (0.002)	0.050*** (0.002)	0.056*** (0.003)
act_ter	-0.083*** (0.003)	-0.072*** (0.004)	-0.071*** (0.003)	-0.063*** (0.003)	-0.057*** (0.002)	-0.068*** (0.004)
formal	0.007*** (0.001)	0.006*** (0.002)	0.002 (0.001)	0.004*** (0.001)	0.010*** (0.001)	0.010*** (0.002)
subocupado	-0.005*** (0.000)	-0.004*** (0.001)	-0.003*** (0.000)	-0.002*** (0.000)	-0.002*** (0.000)	-0.002*** (0.001)
unexplained						
exp_pot	0.149*** (0.028)	0.113*** (0.032)	0.114*** (0.024)	0.087*** (0.022)	0.092*** (0.022)	0.062** (0.031)
exp_pot2	-0.061*** (0.017)	-0.043** (0.020)	-0.027* (0.015)	-0.018 (0.014)	-0.024* (0.014)	0.003 (0.020)
educ2	-0.006 (0.011)	-0.011 (0.013)	-0.008 (0.010)	0.017** (0.009)	-0.007 (0.008)	-0.000 (0.011)
educ3	-0.010 (0.017)	-0.016 (0.022)	-0.011 (0.018)	0.028* (0.015)	-0.017 (0.016)	0.001 (0.021)
educ4	-0.011 (0.012)	-0.015 (0.016)	-0.015 (0.015)	0.019 (0.013)	-0.019 (0.013)	-0.006 (0.018)
educ5	-0.015 (0.014)	-0.016 (0.018)	-0.018 (0.018)	0.016 (0.015)	-0.027* (0.016)	-0.004 (0.020)
educ6	0.000 (0.001)	-0.001 (0.002)	-0.001 (0.002)	0.002 (0.001)	-0.001 (0.002)	0.001 (0.002)

analfabeta	0.001 (0.001)	-0.000 (0.002)	-0.000 (0.001)	0.003** (0.001)	-0.000 (0.001)	0.001 (0.002)
rural	0.024*** (0.003)	0.023*** (0.004)	0.009*** (0.002)	0.013*** (0.003)	0.013*** (0.003)	0.013*** (0.004)
norte	-0.002 (0.002)	-0.008*** (0.003)	-0.010*** (0.002)	-0.010*** (0.002)	-0.003 (0.002)	-0.004 (0.003)
unida	0.111*** (0.006)	0.112*** (0.008)	0.092*** (0.006)	0.093*** (0.005)	0.093*** (0.005)	0.091*** (0.007)
separada	0.008*** (0.002)	0.015*** (0.002)	0.013*** (0.002)	0.012*** (0.002)	0.010*** (0.002)	0.009*** (0.002)
act_sec	0.113*** (0.007)	0.095*** (0.009)	0.113*** (0.007)	0.115*** (0.006)	0.122*** (0.005)	0.101*** (0.008)
act_ter	0.243*** (0.026)	0.172*** (0.033)	0.228*** (0.023)	0.241*** (0.021)	0.268*** (0.019)	0.205*** (0.028)
formal	-0.153*** (0.004)	-0.165*** (0.006)	-0.154*** (0.004)	-0.145*** (0.004)	-0.155*** (0.004)	-0.166*** (0.005)
subocupado	0.002 (0.001)	0.004** (0.002)	0.004*** (0.001)	0.006*** (0.001)	0.004*** (0.001)	0.007*** (0.002)
Constant	-0.039 (0.070)	0.140* (0.085)	0.036 (0.070)	-0.115* (0.061)	0.010 (0.061)	0.072 (0.084)
Observations	181,006	117,450	192,464	213,409	233,872	132,950

Errores estándar por método delta con corrección por conglomerados (UPM).

Datos deduplicados de ENOE. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

De la tabla 3 sobre la descomposición Oaxaca-Heckman del mercado laboral mexicano entre 2018 y 2025, destacan 3 hallazgos clave sobre la estructura de la desigualdad salarial.

En primer lugar, se observa que tanto el grupo de los hombres como de las mujeres muestra un incremento constante en sus ingresos (en escala logarítmica) a lo largo de los seis años, los hombres pasaron de 8.632 en 2018 a 9.155 en 2025, mientras que el grupo de las mujeres subió de 8.344 a 8.816 en el mismo período. No obstante, la diferencia total, muestra una ligera tendencia al alza, iniciando en 0.288 en 2018 y cerrando en 0.340 en 2025.

En segundo lugar, el componente explicado, como se mencionó anteriormente, es negativo en toda la serie. En este sentido, encontramos que la concentración de las mujeres en el sector terciario aporta el componente individual más negativo a la brecha explicada (entre -0.083 en 2018 y -0.068 en 2025). En sentido opuesto, la mayor presencia masculina en el sector secundario contribuye con un signo positivo (entre +0.063 y +0.056). El balance neto entre ambos efectos sectoriales es lo que da forma a una parte sustancial de la brecha explicada total. Este patrón se mantiene estable durante los seis años y es consistente con el diagnóstico de Rosales, Mendoza González y Quintana Romero (2023), quienes documentan la concentración femenina en sectores de servicios con menores primas salariales.

Por su parte, la educación superior (*educ5*) es la variable educativa con mayor impacto negativo (-0.028 en 2018 y -0.018 en 2025). Esto significa que las mujeres tienen una mayor proporción de personas con estudios universitarios, lo que actúa como un fuerte contrapeso para reducir la brecha total. En este sentido, la variable posgrado (*educ6*) opera en la misma dirección, aunque con magnitud menor. Estos signos confirman que las dotaciones de capital humano observables no son la causa de la brecha. En realidad, ayudarían a reducirla si los retornos fueran simétricos.

Los controles geográficos y demográficos desempeñan un papel marginal. La residencia rural y los indicadores de estado civil tienen contribuciones individuales menores a 0.015 log-puntos. La variable norte no resulta significativa en el componente explicado, lo que sugiere que la diferencia en la distribución regional de hombres y mujeres ocupados no contribuye a la brecha.

En tercer lugar, el componente no explicado recoge el grueso de la brecha observada. Encontramos que los coeficientes de experiencias laboral (*exp_pot* y *exp_pot*²) son positivos y altamente significativos (la contribución conjunta oscila entre +0.064 y +0.094 log-puntos según el año). Esto significa que las mujeres reciben menores retornos por cada año adicional de experiencia que los hombres. Esta penalización es uno de los hallazgos más robustos a lo largo del período.

El sector productivo revela que la estructura de remuneración en la economía mexicana castiga desproporcionadamente al grupo de las mujeres. El sector terciario (*act_ter*) es la variable con mayor peso positivo en el componente no explicado, con valores que oscilan entre 0.172 (2019) y 0.268 (2024). Por su parte, el sector secundario (*act_sec*) también contribuye de forma persistente y significativa a la brecha no explicada, con coeficientes estables alrededor de 0.113 en 2018 y 0.101 en 2025. Por lo tanto, el mercado laboral no sólo segrega ocupacionalmente a las mujeres hacia sectores con menores primas, sino que dentro de esos mismos sectores las remunera con retornos menores.

El indicador de unión (*unida*) en el componente no explicado contribuye entre +0.091 y +0.111 log-puntos en todos los años. Esta penalización del matrimonio sobre el salario femenino es consistente con lo que la literatura internacional denomina *motherhood wage penalty* (Olivetti y Petrongolo, 2008) y en el contexto mexicano refleja la división sexual del

trabajo doméstico, donde la formación familiar afecta de manera asimétrica las trayectorias laborales de hombres y mujeres.

La variable formal en el componente no explicado presenta un coeficiente negativo y significativo en todos los años, con magnitudes entre -0.144 y -0.166 log-puntos. Por lo que encontramos que la condición de formalidad reduce la magnitud del componente no explicado de la brecha. Este hallazgo conecta directamente con el análisis sectorial detallado de la sección siguiente. Si la formalidad reduce la penalización femenina, cabe esperar que la brecha total sea sustancialmente diferente entre los segmentos formal e informal del mercado laboral. La sección 4.3 examina formalmente esta hipótesis y constituye, como veremos, el resultado central del estudio.

En conjunto, los factores del componente no explicado indican que la disparidad no se reduce sólo con más educación. El mercado aplica retornos diferenciados a las mismas características en función del sector productivo, del estado civil y del tipo de vínculo laboral. Este último factor, capturado por el signo negativo del coeficiente formal en el componente no explicado, nos indica que la brecha condicional varía sustancialmente entre los segmentos formal e informal del mercado laboral.

4.3 Heterogeneidad por formalidad

La descomposición presentada en 4.1 y 4.2 trata al mercado laboral como un agregado homogéneo. Sin embargo, una característica estructural del mercado mexicano es la coexistencia del empleo formal, sujeto a regulaciones de seguridad social y contractuales, y el empleo informal, donde estas regulaciones están ausentes. Castro Lugo *et al.* (2021) muestran que la convergencia regional de la brecha salarial entre estados es impulsada principalmente por el sector informal; Rosales, Mendoza González y Quintana Romero (2023) documentan que la precariedad laboral amplifica las disparidades de género en el sector turístico.

Las tablas 4 y 5 reportan las descomposiciones Oaxaca-Blinder estimadas por separado para los segmentos formal e informal del mercado laboral. La comparación entre ambas es, a nuestro juicio, el resultado más relevante del estudio.

Tabla 4
Descomposición Oaxaca-Blinder en sector formal

	2018	2019	2022	2023	2024	2025
	b	b	b	b	b	b
overall						
group_1	8.917*** (0.004)	8.968*** (0.004)	9.155*** (0.003)	9.236*** (0.003)	9.368*** (0.003)	9.434*** (0.004)
group_2	8.783*** (0.004)	8.823*** (0.005)	9.020*** (0.004)	9.099*** (0.004)	9.236*** (0.003)	9.315*** (0.005)
difference	0.134*** (0.005)	0.145*** (0.007)	0.136*** (0.005)	0.137*** (0.005)	0.132*** (0.005)	0.118*** (0.006)
explained	-0.059*** (0.003)	-0.051*** (0.004)	-0.046*** (0.003)	-0.045*** (0.002)	-0.044*** (0.002)	-0.052*** (0.003)
unexplained	0.193*** (0.005)	0.196*** (0.006)	0.182*** (0.005)	0.182*** (0.004)	0.176*** (0.004)	0.171*** (0.006)
Observations	43,743	29,817	47,416	53,197	57,902	29,960

Errores estándar robustos. Datos deduplicados de ENOE.

*** p<0.01, ** p<0.05, * p<0.10.

Tabla 5. Descomposición Oaxaca-Blinder en sector informal

	2018	2019	2022	2023	2024	2025
	b	b	b	b	b	b
overall						
group_1	8.369*** (0.004)	8.433*** (0.005)	8.645*** (0.004)	8.708*** (0.004)	8.830*** (0.004)	8.895*** (0.005)
group_2	7.952*** (0.006)	7.978*** (0.007)	8.219*** (0.005)	8.290*** (0.005)	8.404*** (0.005)	8.446*** (0.007)
difference	0.417*** (0.007)	0.454*** (0.009)	0.426*** (0.007)	0.418*** (0.007)	0.426*** (0.006)	0.449*** (0.009)
explained	-0.106*** (0.004)	-0.092*** (0.005)	-0.068*** (0.004)	-0.071*** (0.004)	-0.067*** (0.004)	-0.081*** (0.005)
unexplained	0.523*** (0.007)	0.546*** (0.009)	0.494*** (0.007)	0.489*** (0.007)	0.493*** (0.007)	0.530*** (0.009)
Observations	48,685	31,769	49,990	56,482	60,177	33,316

Errores estándar robustos. Datos deduplicados de ENOE.

*** p<0.01, ** p<0.05, * p<0.10.

La diferencia salarial entre hombres y mujeres dentro del sector formal oscila entre 0.118 y 0.145 log-puntos durante el período, con un valor promedio de 0.134. Esto representa el 13% de diferencia condicional. Por su parte, en el sector informal, la diferencia oscila entre 0.417 y 0.454, con un valor promedio de 0.432, cerca del 43%. La *ratio* informal/formal de la brecha es de 3.1 a 3.8 veces a lo largo del período. La tabla 6 sintetiza esta comparación para facilitar la lectura:

Tabla 6
Ratio informal/formal de la brecha salarial

Año	Brecha formal	Brecha informal	Ratio
2018	0.134***	0.417***	3.1
2019	0.145***	0.454***	3.1
2022	0.136***	0.426***	3.1
2023	0.137***	0.418***	3.0
2024	0.132***	0.426***	3.2
2025	0.118***	0.449***	3.8

Fuente: elaboración propia con datos de ENOE (INEGI, 2018-2025), datos deduplicados. Estimaciones de las tablas 4 y 5. Brecha medida en logaritmo del ingreso mensual; ratio = brecha informal / brecha formal. Errores estándar robustos con cluster por UPM. *** $p<0.01$, ** $p<0.05$, * $p<0.10$.

Este hallazgo es robusto, estable en el tiempo y de magnitud económica considerable. La brecha global del orden del 30%, documentada en la sección 4.1, resulta ser un promedio que oculta dos realidades muy distintas: por un lado, una brecha relativamente contenida en el sector formal y, por el otro lado, una brecha sustancialmente mayor en el sector informal.

En el componente explicado, ambos sectores muestran valores negativos consistentes con lo discutido en 4.2 sobre que el promedio de las dotaciones observables de las mujeres es mejor que las de los hombres en ambos segmentos, pero con diferente magnitud. Por ejemplo, en el sector formal el componente explicado oscila alrededor de -0.050 log-puntos, mientras que en el sector informal alcanza valores cercanos a -0.080 , esto sugiere que las mujeres en el sector informal presentan, en términos relativos, una composición más favorable que los hombres informales que la que se observa entre formales.

En el componente no explicado la diferencia entre sectores es mayor: el sector formal promedia 0.184 log-puntos, mientras que en el sector informal promedia 0.512 log-puntos. Si la brecha se debiera a variables omitidas como productividad latente, esfuerzo o preferencias laborales, no habría razón de que estos sesgos fueran 2.8 veces más severos en el sector informal que en el formal, ya que las mismas variables omitidas operarían de manera similar en ambos segmentos.

Aunque ninguno de estos mecanismos elimina por completo la brecha (siguen existiendo 13 puntos porcentuales de diferencia condicional en el sector formal), sí parecen contener su magnitud. En el sector informal, donde estos mecanismos están ausentes, la brecha no explicada se

expande hasta valores que, dado el contraste con el sector formal, resulta razonable interpretar como evidencia más sólida de discriminación salarial estructural.

Esta lectura es congruente con el marco analítico desarrollado por Castro Lugo *et al.* (2021), quienes argumentan que la reducción de la brecha nacional ha estado impulsada principalmente por dinámicas dentro del sector informal, y con la hipótesis de Rosales *et al.* (2023) según la cual la precariedad laboral amplifica las desigualdades de género al desactivar las instituciones que las contendrían.

Esta heterogeneidad sectorial reordena la interpretación de los hallazgos de 4.1 y 4.2. El componente no explicado promedio de 0.372 que se documentó para el conjunto del mercado laboral es una mezcla ponderada entre dos realidades muy distintas: por un lado, un sector formal con penalizaciones moderadas y, por el otro, un sector informal donde el grueso de la disparidad de género se concentra.

4.4 Estabilidad temporal y modelo de selección

La descomposición presentada en las secciones anteriores se estima año por año, así ofrece una visión transversal repetida, pero no responde formalmente dos preguntas que conciernen directamente al diseño del estudio: ¿la brecha salarial cambió significativamente entre los años analizados? y ¿la decisión de participar en el mercado laboral introduce un sesgo de selección que distorsione las estimaciones? A continuación, abordamos ambas mediante un test formal de igualdad de la brecha entre años y la inspección de los parámetros del modelo de selección de Heckman.

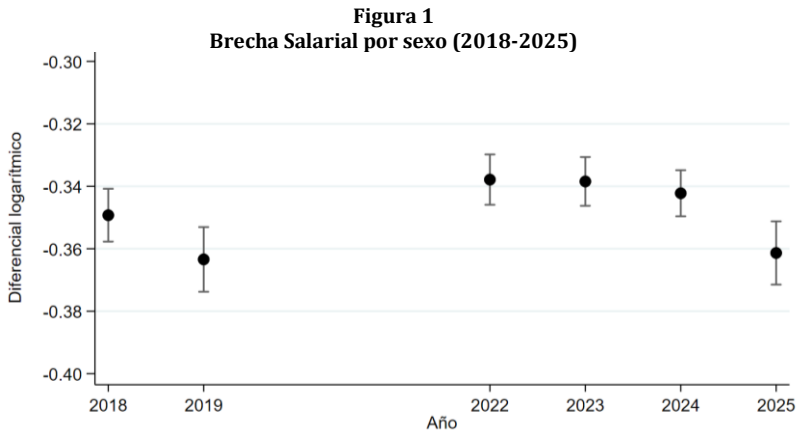
Para evaluar si la brecha salarial se modificó durante 2018-2025, se estima un modelo apilado de salarios con interacciones sexo × año, tomando a los hombres de 2018 como categoría de referencia. La tabla 7 reporta los coeficientes relevantes; la figura 1 visualiza la evolución del efecto marginal de ser mujer en cada año.

Tabla 7
Test formal de diferencias en la brecha de género entre años

	lsalario	se
mujer=1	-0.349***	(0.004)
mujer=1 x ano=2019	-0.014**	(0.007)
mujer=1 x ano=2022	0.011*	(0.006)

mujer=1 x ano=2023	0.011*	(0.006)
mujer=1 x ano=2024	0.007	(0.006)
mujer=1 x ano=2025	-0.012*	(0.007)
Observaciones	542,454	
R ²	0.372	

F(5, 61643) = 5.41, p = 0.0001 (test de igualdad de interacciones).
Referencia: hombres en 2018. N = 542,454 observaciones.
Errores estándar agrupados a nivel UPM.



Fuente: cálculos del autor, con base en ENOE (INEGI, 2018-2025), datos deduplicados.
Modelo apilado con controles individuales completos.
Errores estándar agrupados a nivel UPM, IC al 95%.

El test conjunto sobre las cinco interacciones sexo $\times$ año rechaza la hipótesis nula de igualdad de la brecha entre años con un estadístico $F(5, 61643) = 5.41$ y $p = 0.0001$. Formalmente, la brecha no es idéntica a lo largo del período, sin embargo, sus magnitudes oscilan entre -0.338 (en 2022 y 2023) y -0.363 (en 2019), un rango de apenas 2.6 puntos porcentuales sobre una brecha promedio cercana al 34%. Las interacciones individuales sexo $\times$ año tienen magnitudes entre -0.014 y $+0.011$, valores que, aun siendo estadísticamente significativos en algunos casos por el tamaño de la muestra, son sustantivamente pequeños.

Observamos que la brecha más amplia ocurre en 2019, justo antes de la pandemia, y vuelve a ampliarse en 2025. Por su parte, los años 2022-2024 muestran brechas relativamente más estrechas. Esto se debe a que, durante la recuperación postpandemia, la precariedad laboral masculina se intensificó. Lo que implica que la brecha femenina se redujo no por una mejora, sino por un deterioro relativo de la posición de los hombres. Por lo tanto, no encontramos evidencia de una ruptura estructural atribuible a la pandemia. Esta hipótesis es consistente con el argumento de Rosales

et al. (2023), quienes documentan que la brecha de género tiende a estrecharse cuando la precariedad masculina aumenta.

La estimación del modelo Heckman permite evaluar si el sesgo de selección muestral afecta las conclusiones de la descomposición. La tabla 8 reporta los tres parámetros relevantes para los seis años: ρ (correlación entre los errores de la ecuación de selección y la ecuación salarial), σ (desviación estándar del error salarial) y λ (inverso de la *ratio* de Mills, igual al producto $\rho \cdot \sigma$).

Tabla 8
Parámetros del modelo de selección Heckman para la ecuación salarial femenina

Año	ρ	σ	λ	LR_chi2	p-LR	N
2018	-0.010	0.631	-0.006	0.11	[0.743]	125,508
2019	0.038	0.642	0.025	2.61	[0.106]	80,394
2022	0.037	0.625	0.023*	4.36	[0.037]	135,620
2023	0.044	0.631	0.028***	8.31	[0.004]	148,218
2024	0.036	0.634	0.023***	9.22	[0.002]	165,924
2025	0.042	0.646	0.027***	4.91	[0.027]	96,431

Fuente: elaboración propia con datos de ENOE (INEGI, 2018-2025), datos deduplicados. Estimaciones Heckman MLE con *tiene_hijos* e *hijos_3mas* como restricciones de exclusión. $\rho = \tanh(\text{athrho})$; $\sigma = \exp(\text{lnsigma})$; $\lambda = \rho \cdot \sigma$. LR_chi² evalúa $H_0: \rho = 0$; rechazo ($p < 0.05$) indica sesgo de selección significativo. Errores estándar robustos con cluster por UPM.

Los resultados muestran que en 2018 y 2019, la correlación entre los errores de selección y salario no es estadísticamente distinguible de cero ($p = 0.743$ y $p = 0.106$, respectivamente), pero en los años 2022, 2023, 2024 y 2025, el test de razón de verosimilitud rechaza la hipótesis nula de independencia ($p < 0.05$ en todos los casos), lo que indica que en la segunda mitad del período existe una correlación detectable entre los procesos de selección laboral y salario. No obstante, las magnitudes de la correlación ρ son uniformemente pequeñas (entre -0.010 y $+0.044$), y los valores del coeficiente λ se sitúan entre -0.006 y $+0.028$. Aunque la corrección de Heckman captura un sesgo estadísticamente significativo en cuatro de los seis años, sin embargo, su magnitud es insuficiente para alterar las conclusiones de la descomposición Oaxaca-Blinder.

Por lo tanto, las conclusiones presentadas en las secciones 4.1 a 4.3 sobre la magnitud de la brecha son robustas a la corrección por selección muestral. No obstante, los valores del parámetro λ transitan de signo negativo en 2018 a signo positivo en 2025, con el mismo movimiento

observable en p . Esta evolución nos indica que, si en 2018 las mujeres ocupadas tendían a estar negativamente seleccionadas respecto a sus características inobservables, hacia 2025 esta selección parece haberse invertido hacia un patrón positivo, dicho cambio observable, no es un hallazgo estadísticamente robusto.

La siguiente subsección complementa esta agenda de robustez con un ejercicio de sensibilidad al grupo de referencia en la descomposición.

4.5 Robustez al grupo de referencia

A continuación, se presenta un ejercicio de robustez en el que la descomposición se reestima utilizando los coeficientes masculinos como referencia, en la formulación clásica de Oaxaca (1973). La tabla 9 reporta los resultados de esta especificación alternativa, descompuesta esta vez en tres términos: dotaciones, coeficientes e interacción.

Tabla 9
Descomposición Oaxaca-Blinder clásica con coeficientes masculinos

	2018 b	2019 b	2022 b	2023 b	2024 b	2025 b
overall						
group_1	8.632*** (0.003)	8.695*** (0.004)	8.894*** (0.003)	8.966*** (0.003)	9.099*** (0.003)	9.155*** (0.004)
group_2	8.337*** (0.004)	8.381*** (0.005)	8.607*** (0.004)	8.678*** (0.004)	8.802*** (0.004)	8.846*** (0.005)
difference	0.295*** (0.005)	0.314*** (0.006)	0.288*** (0.005)	0.288*** (0.005)	0.297*** (0.004)	0.309*** (0.006)
	-	-	-	-	-	-
endowments	0.082*** (0.005)	0.076*** (0.006)	0.065*** (0.004)	0.061*** (0.004)	0.043*** (0.004)	0.056*** (0.005)
coefficients	0.353*** (0.005)	0.369*** (0.006)	0.336*** (0.004)	0.333*** (0.004)	0.334*** (0.004)	0.355*** (0.006)
interaction	0.024*** (0.004)	0.021*** (0.005)	0.017*** (0.003)	0.015*** (0.003)	0.007** (0.003)	0.010** (0.004)
Observations	92,428	61,586	97,406	109,679	118,079	63,276

Componentes: dotaciones, coeficientes e interacción.

Errores estándar robustos. Datos deduplicados de ENOE.

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

La comparación entre las tablas 2 (especificación principal) y 9 (especificación alternativa) confirma que la estructura cualitativa de la brecha salarial no depende de la elección del grupo de referencia. Cabe aclarar que la tabla 9 se refiere únicamente a la población ocupada (hombres + mujeres ocupadas de la tabla 1).

Encontramos que el componente de dotaciones es negativo en los seis años, con magnitudes entre -0.043 y -0.082 log-puntos, ligeramente menores que las del componente explicado de la tabla 2 (que oscilan entre -0.037 y -0.066), pero comparten el mismo signo y orden de magnitud, es decir, las mujeres ocupadas tienen, en promedio, mejores dotaciones observables que los hombres ocupados. Por otro lado, el componente de coeficientes es positivo y significativo en todos los años, con magnitudes entre 0.333 y 0.369 , también resulta consistente con el componente no explicado de la tabla 2, que oscila entre 0.354 y 0.399 . Las diferencias entre ambas especificaciones, de 0.02 a 0.03 log-puntos, reflejan el cambio en el vector de referencia y la ausencia de corrección de selección en esta especificación alternativa. Por último, el componente de interacción captura el efecto conjunto de las diferencias en dotaciones y en cómo se valoran, oscilan entre 0.007 y 0.024 , su magnitud reducida indica que la mayor parte de la variación se captura adecuadamente con los dos componentes principales y que la interpretación basada en una descomposición doble no introduce distorsiones sustanciales respecto a la triple descomposición.

En conjunto, los resultados presentados en esta sección de robustez confirman la solidez de los hallazgos principales. La brecha salarial de género en México durante 2018-2025 es un fenómeno estable de magnitud aproximada del 30%, cuyo componente no atribuible a dotaciones observables se concentra desproporcionadamente en el sector informal, y cuyos resultados no dependen ni del grupo de referencia elegido ni de la corrección por selección.

Conclusiones

Los resultados confirman que la brecha salarial entre hombres y mujeres es un fenómeno estructural y persistente, mediante la descomposición Oaxaca-Blinder con corrección de selección de Heckman sobre datos deduplicados de la ENOE encontramos que la brecha total se mantuvo en torno al 30% durante los seis años analizados, con oscilaciones de 2.6 puntos porcentuales, la disparidad sugiere que las fuerzas que generan la desigualdad están profundamente arraigadas en el sistema laboral mexicano, más allá de los ciclos de corto plazo y de las políticas implementadas durante el período de estudio.

El capital humano observable es condición necesaria, pero insuficiente para reducir la brecha. Como se documentó, las mujeres ocupadas en México han superado a los hombres en escolaridad media y en la

proporción de quienes alcanzan niveles educativos superiores; sin embargo, esta ventaja en dotaciones no se traduce en paridad salarial. Por su parte, el componente explicado de la descomposición es negativo en los seis años, lo que indica que, si los retornos a las características observables fueran simétricos entre hombres y mujeres, la brecha se invertiría en favor de las mujeres. Una posible explicación de la desigualdad es cómo el mercado laboral valora y remunera las mismas características.

Las contribuciones combinadas de los términos lineal y cuadrático de la experiencia potencial al componente no explicado se sitúan entre 0.064 y 0.094 log-puntos a lo largo del período. Significa que las mujeres reciben menos por cada año adicional de experiencia que los hombres, por lo que la disparidad se reproduce en cohortes sucesivas en lugar de diluirse con el tiempo. El hallazgo es consistente en los seis años analizados y robusto a la corrección por selección muestral.

El sector informal constituye el núcleo de la disparidad salarial de género en México. El mercado laboral mexicano oculta dos realidades: una brecha del 13% en el sector formal y una brecha del 43% en el sector informal. La *ratio* informal/formal oscila entre 3.0 y 3.8 a lo largo del período, sin alteración cualitativa. La descomposición Oaxaca-Blinder estimada por separado para cada segmento revela que el componente no atribuible a dotaciones es 2.8 veces mayor en el sector informal que en el formal, patrón que se mantiene durante los seis años analizados.

Además, las mujeres ocupadas en empleo informal tienen dotaciones de capital humano aún más favorables respecto a los hombres (efecto de -0.080) que en el sector formal (-0.050), es decir, la mujer está sobrecalificada en comparación al hombre informal en una magnitud que excede a la sobrecalificación femenina dentro del sector formal, y aun así enfrenta una penalización salarial residual sustancialmente mayor.

La formalización del empleo aparece con potencial igualador salarial debido a los contratos colectivos, escalas salariales formalizadas, sindicalización y supervisión institucional del cumplimiento laboral. No obstante, la formalización no implica que sea una solución, ya que la brecha persiste alrededor del 13% en el sector formal. A pesar de ello, la expansión de la cobertura del empleo formal constituye un mecanismo complementario para la reducción de la disparidad salarial. Por su parte, la subocupación opera como un mecanismo adicional de precariedad laboral diferenciada por género.

El sesgo de selección laboral femenino de Heckman con restricciones de exclusión válidas (*tiene_hijos e hijos_3mas*) es estadísticamente significativo en cuatro de los seis años analizados (2022-2025). La magnitud y los factores determinantes de la brecha son robustos a la corrección por selección. Aun así, el parámetro de selección transita de signo negativo en 2018 a positivo en 2025, sugiriendo una posible transformación post pandemia. No obstante, este patrón carece de la significancia estadística necesaria para sostener una afirmación causal.

Finalmente, el componente no explicado revela factores fundamentales que actúan como barreras invisibles en la trayectoria de las mujeres. En primer lugar, el estado de unión conyugal opera como una penalización estructural que suma sistemáticamente cerca de 0.10 log-puntos a la disparidad, reflejando el peso de los roles domésticos en la valoración salarial. En segundo lugar, el sector terciario aporta los coeficientes más elevados a la parte no explicada (hasta 0.268), lo que confirma que la segregación hacia los servicios no sólo es ocupacional, sino que conlleva una profunda inequidad en la remuneración de las capacidades dentro de este sector.

Declaraciones

1. *Financiamiento*

La presente investigación no recibió ninguna subvención ni apoyo económico específico de agencias de financiamiento del sector público, del sector comercial o de agencias sin fines de lucro.

2. *Contribución de autoría*

La atribución de roles de autoría para este manuscrito se define de la siguiente manera:

[J.A.P.G.]: Conceptualización, Metodología, Software, Validación, Análisis formal, Investigación, Recursos, Curación de datos, Redacción – borrador original, Redacción – revisión y edición, Visualización, Supervisión, Administración del proyecto.

[S.A.H.]: Conceptualización, Metodología, Software, Validación, Análisis formal, Investigación, Recursos, Curación de datos, Redacción – borrador original, Redacción – revisión y edición, Visualización, Supervisión, Administración del proyecto.

Todos los autores han leído y aceptado la versión final del manuscrito.

3. Conflicto de interés

Los autores declaran que no tienen intereses financieros compitiendo, ni relaciones personales directas o indirectas que pudieran haber influido o aparentar haber influido en la realización, los resultados o la interpretación de la investigación presentada en este manuscrito.

4. Disponibilidad de datos

Los datos que respaldan los resultados de esta investigación están disponibles a través del autor de correspondencia previa solicitud razonable.

5. Declaración de uso de Inteligencia Artificial Generativa (IAG)

Los autores declaran no haber utilizado herramientas de Inteligencia Artificial generativa ni tecnologías asistidas por IA para la preparación, análisis de datos o redacción de este manuscrito.

Referencias

- [1] Arceo-Gómez, E. O., & Campos-Vázquez, R. M. (2014). Evolución de la brecha salarial de género en México. *El Trimestre Económico*, 81(323), 619–653. <https://doi.org/10.20430/ete.v81i323.125>
- [2] Arceo Gómez, E. O., Campos Vázquez, R. M., Medina Cortina, E. M., & Vélez Grajales, R. (2018). Negociación y preferencias económicas por género: evidencia experimental en México. *El Trimestre Económico*, 85(339), 645–678. <https://doi.org/10.20430/ete.v85i339.523>
- [3] Arceo-Gómez, E. O., & Torres López, P. J. (2021). Brechas salariales por autoidentificación indígena y rasgos lingüísticos en México. *Sobre México. Temas de Economía. Nueva Época*, 1(3), 129–161. <https://doi.org/10.48102/rsm.vi3.94>
- [4] Becker, G. S. (1957). *The economics of discrimination*. University of Chicago Press.
- [5] Blinder, A. S. (1973). Wage discrimination: Reduced form and structural estimates. *Journal of Human Resources*, 8(4), 436–455. <https://doi.org/10.2307/144855>
- [6] Bonavida, C., & Gasparini, L. (2024). *El rol de las habilidades y las tareas en las brechas de género en América Latina* (Documento de Trabajo del CEDLAS N.º 330). Centro de Estudios Distributivos, Laborales y Sociales, Universidad Nacional de La Plata. [doc.cedlas330.pdf](https://cedlas.org/argentina/doc/cedlas330.pdf)
- [7] Cameron, A. C., & Miller, D. L. (2015). A practitioner's guide to cluster-robust inference. *Journal of Human Resources*, 50(2), 317–372. <https://doi.org/10.3368/jhr.50.2.317>
- [8] Canedo, A. (2019). Labor market discrimination against indigenous peoples in Mexico: A decomposition analysis of wage differentials. *Iberoamericana—Nordic Journal of Latin American and Caribbean Studies*, 48(1), 12–27. <https://doi.org/10.16993/iberoamericana.433>
- [9] Castro Lugo, D., Rodríguez Pérez, R. E., & Ramos Lobo, R. (2021). ¿Existe convergencia regional en la brecha salarial por género en México? *Economía*,

- Sociedad y Territorio*, 21(67), 685–717.
<https://doi.org/10.22136/est20211633>
- [10] Choe, C., Jung, S., & Oaxaca, R. L. (2022). What's the risk from competing? Competition aversion and the gender wage gap (IZA Discussion Paper No. 15048). IZA Institute of Labor Economics.
<https://www.iza.org/publications/dp/15048/whats-the-risk-from-competing-competition-aversion-and-the-gender-wage-gap>
- [11] Cuéllar, C. Y., & Moreno, J. O. (2022). Employment, wages, and the gender gap in Mexico: Evidence of three decades of the urban labor market. *Latin American Journal of Central Banking*, 3(2), 100055.
<https://doi.org/10.1016/j.latch.2022.100055>
- [12] De la Luz Tovar, C., & Pineda Salazar, S. I. (2023). Análisis regional de la discriminación salarial de los trabajadores indígenas en México. *Nóesis. Revista de Ciencias Sociales y Humanidades*, 32(64), 47–72.
<https://doi.org/10.20983/noesis.2023.2.3>
- [13] Fortin, N. M. (2008). The gender wage gap among young adults in the United States: The importance of money versus people. *Journal of Human Resources*, 43(4), 884–918. <https://doi.org/10.3368/jhr.43.4.884>
- [14] Heckman, J. J. (1979). Sample selection bias as a specification error. *Econometrica*, 47(1), 153–161. <https://doi.org/10.2307/1912352>
- [15] Instituto Nacional de Estadística y Geografía. (2018–2019, 2022–2025). *Encuesta Nacional de Ocupación y Empleo (ENOE), población de 15 años y más de edad: microdatos* [Conjunto de datos].
<https://www.inegi.org.mx/programas/enoe/15ymas/>
- [16] Jann, B. (2008). The Blinder–Oaxaca decomposition for linear regression models. *The Stata Journal*, 8(4), 453–479.
<https://doi.org/10.1177/1536867X0800800401>
- [17] Jung, S. (2014). The gender wage gap and sample selection via risk attitudes (THEMA Working Paper No. 2014-25). Université de Cergy-Pontoise.
<https://shs.hal.science/halshs-00965520v2>
- [18] Kröger, H., & Hartmann, J. (2021). Extending the Kitagawa–Oaxaca–Blinder decomposition approach to panel data. *The Stata Journal*, 21(2), 360–410.
<https://doi.org/10.1177/1536867X211025800>
- [19] Linthon Delgado, D. E., & Méndez-Heras, L. B. (2022). Descomposición de la brecha salarial de género en el Ecuador. *Revista Mexicana de Economía y Finanzas Nueva Época*, 17(1), e706.
<https://doi.org/10.21919/remef.v17i1.706>
- [20] Mincer, J. (1974). Schooling, experience, and earnings. National Bureau of Economic Research. <https://www.nber.org/books-and-chapters/schooling-experience-and-earnings>
- [21] Mussa, R. (2014). Extending the Oaxaca–Blinder decomposition to the independent double hurdle model: With application to parental spending on education in Malawi. *Studies in Economics and Econometrics*, 38(3), 39–54.
<https://doi.org/10.1080/10800379.2014.12097271>
- [22] Mulligan, C. B., & Rubinstein, Y. (2008). Selection, investment, and women's relative wages over time. *The Quarterly Journal of Economics*, 123(3), 1061–1110. <https://doi.org/10.1162/qjec.2008.123.3.1061>

- [23] Oaxaca, R. (1973). Male-female wage differentials in urban labor markets. *International Economic Review*, 14(3), 693–709. <https://doi.org/10.2307/2525981>
- [24] Oaxaca, R. L., & Ransom, M. R. (1994). On discrimination and the decomposition of wage differentials. *Journal of Econometrics*, 61(1), 5–21. [https://doi.org/10.1016/0304-4076\(94\)90074-4](https://doi.org/10.1016/0304-4076(94)90074-4)
- [25] Oaxaca, R. L., & Sierminska, E. (2023). *Oaxaca–Blinder meets Kitagawa: What is the link?* (IZA Discussion Paper No. 16188). IZA Institute of Labor Economics. <https://www.iza.org/publications/dp/16188/oaxaca-blinder-meets-kitagawa-what-is-the-link>
- [26] Olivetti, C., & Petrongolo, B. (2008). Unequal pay or unequal employment? A cross-country analysis of gender gaps. *Journal of Labor Economics*, 26(4), 621–654. <https://doi.org/10.1086/589458>
- [27] Ospino, C. G., Roldán Vásquez, P., & Barraza Narváez, N. (2010). Oaxaca–Blinder wage decomposition: Methods, critiques and applications. A literature review [La descomposición salarial de Oaxaca–Blinder: métodos, críticas y aplicaciones. Una revisión de la literatura]. *Revista de Economía del Caribe*, (5), 237–274. <https://doi.org/10.14482/ecoca.05.909.864>
- [28] Otsu, T., & Tanaka, S. (2022). Empirical likelihood inference for Oaxaca–Blinder decomposition. *Economics Letters*, 219, 110812. <https://doi.org/10.1016/j.econlet.2022.110812>
- [29] Phelps, E. S. (1972). The statistical theory of racism and sexism. *American Economic Review*, 62(4), 659–661. <http://www.jstor.org/stable/1806107>
- [30] Polachek, S. W. (1981). Occupational self-selection: A human capital approach to sex differences in occupational structure. *The Review of Economics and Statistics*, 63(1), 60–69. <https://doi.org/10.2307/1924218>
- [31] Popli, G. K. (2013). Gender wage differentials in Mexico: A distributional approach. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 176(2), 295–319. <https://doi.org/10.1111/j.1467-985X.2011.01031.x>
- [32] Puhani, P. A. (2000). The Heckman correction for sample selection and its critique. *Journal of Economic Surveys*, 14(1), 53–68. <https://doi.org/10.1111/1467-6419.00104>
- [33] Rahimi, E., & Hashemi Nazari, S. S. (2021). A detailed explanation and graphical representation of the Blinder–Oaxaca decomposition method with its application in health inequalities. *Emerging Themes in Epidemiology*, 18, Article 12. <https://doi.org/10.1186/s12982-021-00100-9>
- [34] Rodríguez Pérez, R. E. (2018). Brecha salarial por género en México: desde un enfoque regional, según su exposición a la apertura comercial 2005–2015. *Nóesis. Revista de Ciencias Sociales y Humanidades*, 27(54), 19–38. <https://doi.org/10.20983/noesis.2018.2.2>
- [35] Rodríguez Pérez, R. E. (2019). Diferencial salarial por género entre el sector público y privado formal-informal en México. *Revista de Economía*, 36(93), 62–89. <https://doi.org/10.33937/reveco.2019.108>
- [36] Romero Martínez, V. M., & Prudencio Vázquez, J. A. (2022). Brecha salarial de género: análisis exploratorio con la ENOE, 2013 y 2018. En M. del P. A. Mora Cantellano, S. E. Serrano Oswald, & V. E. Mota Flores (Coords.), *Reconfigurando territorios a partir de la cultura, el empoderamiento de las mujeres y nuevos turismo* (Vol. 4, pp. 299–318). Universidad Nacional

- Autónoma de México; Asociación Mexicana de Ciencias para el Desarrollo Regional. <http://ru.iiec.unam.mx/5895/>
- [37] Rosales, R. A., Mendoza González, M. Á., & Quintana Romero, L. (2023). Análisis espacial de la desigualdad salarial y la precariedad laboral en el sector turístico mexicano. *Contaduría y Administración*, 68(2), e392. <https://doi.org/10.22201/fca.24488410e.2023.4828>
- [38] Salce Díaz, F. (2021). Evolución y análisis de la discriminación salarial por género en Chile. *El Trimestre Económico*, 88(349), 39–75. <https://doi.org/10.20430/ete.v88i349.984>

Anexo A. Estadísticos descriptivos por año y sexo

Tabla A1
Medias por año y sexo, ENOE 2018-2025 (datos deduplicados)

	(1) H18 mean	(2) M18 mean	(3) H19 mean	(4) M19 mean	(5) H22 mean	(6) M22 mean	(7) H23 mean
Edad	38.462	38.756	38.827	39.380	38.769	38.803	39.044
Años de escolaridad	9.551	10.236	9.688	10.210	10.039	10.715	10.082
exp_pot	23.009	22.628	23.199	23.243	22.808	22.142	23.062
Número de hijos	.	2.041	.	2.080	.	1.896	.
salario	6773.16	5282.74	7157.47	5525.75	8712.11	6863.34	9459.97
Isalario	6	2	7	1	5	5	5
rsalario	8.567	8.278	8.630	8.310	8.832	8.547	8.907
rural	0.621	0.559	0.582	0.521	0.354	0.325	0.574
Estado							
fronterizo	0.196	0.202	0.200	0.218	0.209	0.213	0.210
del norte							
formal	0.436	0.427	0.436	0.432	0.442	0.446	0.442
subocupado	0.082	0.059	0.083	0.071	0.091	0.082	0.087
unida	0.675	0.518	0.669	0.526	0.639	0.499	0.633
separada	0.055	0.167	0.061	0.168	0.064	0.174	0.066
soltera	0.270	0.316	0.271	0.306	0.297	0.327	0.301
tiene_hijos	.	0.748	.	0.758	.	0.728	.
hijos_3mas	.	0.351	.	0.360	.	0.321	.
educ1	0.035	0.033	0.033	0.034	0.027	0.024	0.028
educ2	0.234	0.202	0.223	0.199	0.204	0.164	0.203
educ3	0.335	0.297	0.330	0.296	0.322	0.291	0.322
educ4	0.224	0.204	0.229	0.203	0.241	0.233	0.239
educ5	0.162	0.247	0.173	0.248	0.192	0.264	0.192
educ6	0.010	0.017	0.011	0.018	0.014	0.023	0.015
analfabeta	0.030	0.030	0.028	0.028	0.025	0.021	0.026
act_sec	0.342	0.192	0.336	0.186	0.342	0.192	0.331
act_ter	0.490	0.781	0.504	0.785	0.504	0.774	0.511
Observations	55,498	36,930	37,056	24,530	56,844	40,562	65,191

Tabla A1
Medias por año y sexo, ENOE 2018-2025 (datos deduplicados) Continuación

	(8) M23 mean	(9) H24 mean	(10) M24 mean	(11) H25 mean	(12) M25 mean
Edad	39.432	39.556	39.694	39.709	40.011
Años de escolaridad	10.743	10.182	10.937	10.286	11.029
exp_pot	22.775	23.437	22.838	23.586	23.110
Número de hijos	1.934	.	1.867	.	1.888
salario	7371.239	10835.921	8496.081	11624.524	8937.708
lsalario	8.613	9.048	8.752	9.121	8.800
rural	0.522	0.575	0.518	0.596	0.534
Estado fronterizo del norte	0.211	0.204	0.212	0.198	0.204
formal	0.434	0.459	0.439	0.451	0.426
subocupado	0.082	0.083	0.078	0.081	0.074
unida	0.494	0.629	0.496	0.646	0.502
separada	0.179	0.069	0.177	0.067	0.189
soltera	0.327	0.302	0.327	0.287	0.309
tiene_hijos	0.729	.	0.723	.	0.737
hijos_3mas	0.328	.	0.317	.	0.326
educ1	0.028	0.027	0.024	0.027	0.024
educ2	0.164	0.196	0.152	0.197	0.153
educ3	0.289	0.313	0.285	0.309	0.286
educ4	0.227	0.244	0.237	0.256	0.233
educ5	0.270	0.204	0.274	0.196	0.275
educ6	0.021	0.016	0.027	0.014	0.027
analfabeta	0.026	0.024	0.021	0.025	0.022
act_sec	0.185	0.331	0.179	0.331	0.177
act_ter	0.780	0.520	0.789	0.514	0.789
Observations	44,488	67,948	50,131	36,519	26,757

Fuente: elaboración propia con datos de la ENOE (INEGI, 2018-2025), datos deduplicados. Cifras reportadas como medias muestrales con desviaciones estándar entre paréntesis. La submuestra de hombres ocupados incluye a hombres en edad laboral con ingreso positivo; la de mujeres ocupadas, a mujeres en edad laboral con ingreso positivo; la de mujeres no ocupadas, a mujeres en edad laboral sin ingreso laboral observado durante el trimestre de referencia.

Anexo B
Estimación Probit de la ecuación de selección laboral femenina.

	(1) 2018 b	(2) 2019 b	(3) 2022 b	(4) 2023 b	(5) 2024 b	(6) 2025 b
ocupado						
exp_pot	0.046*** (0.001)	0.047*** (0.001)	0.046*** (0.001)	0.044*** (0.001)	0.045*** (0.001)	0.045*** (0.001)
exp_pot2	-0.092*** (0.002)	-0.094*** (0.002)	-0.096*** (0.002)	-0.095*** (0.002)	-0.095*** (0.001)	-0.096*** (0.002)
educ2	-0.112** (0.048)	-0.298*** (0.060)	-0.257*** (0.049)	-0.313*** (0.046)	-0.280*** (0.046)	-0.208*** (0.058)
educ3	-0.223*** (0.049)	-0.412*** (0.061)	-0.350*** (0.050)	-0.449*** (0.046)	-0.403*** (0.047)	-0.317*** (0.059)
educ4	-0.114** (0.049)	-0.319*** (0.062)	-0.255*** (0.050)	-0.367*** (0.046)	-0.314*** (0.047)	-0.237*** (0.059)
educ5	-0.001 (0.049)	-0.230*** (0.061)	-0.165*** (0.050)	-0.268*** (0.046)	-0.208*** (0.047)	-0.161*** (0.059)
educ6	0.195*** (0.058)	-0.054 (0.072)	0.053 (0.056)	-0.095* (0.052)	-0.004 (0.052)	0.036 (0.066)
analfabeta	-0.025 (0.051)	-0.160** (0.065)	-0.155*** (0.053)	-0.140*** (0.049)	-0.160*** (0.049)	-0.027 (0.062)
rural	-0.007 (0.009)	-0.051*** (0.011)	-0.014 (0.010)	-0.012 (0.008)	0.011 (0.008)	0.014 (0.010)
Estado fronterizo	0.019* (0.010)	0.085*** (0.012)	0.066*** (0.009)	0.078*** (0.009)	0.052*** (0.008)	0.007 (0.011)
del norte	-0.280*** (0.013)	-0.253*** (0.016)	-0.185*** (0.012)	-0.178*** (0.012)	-0.155*** (0.011)	-0.154*** (0.015)
unida	0.205*** (0.017)	0.212*** (0.021)	0.224*** (0.016)	0.232*** (0.015)	0.226*** (0.014)	0.242*** (0.018)
separada	0.242*** (0.014)	0.234*** (0.018)	0.214*** (0.014)	0.194*** (0.013)	0.166*** (0.012)	0.179*** (0.016)
tiene_hijos	-0.047*** (0.010)	-0.032** (0.013)	-0.040*** (0.010)	-0.031*** (0.009)	-0.025*** (0.009)	-0.010 (0.012)
hijos_3mas	-0.729*** (0.049)	-0.509*** (0.062)	-0.575*** (0.050)	-0.458*** (0.046)	-0.506*** (0.047)	-0.665*** (0.059)
Constant						
Observaciones	125,508	80,394	135,620	148,218	165,924	96,431
Pseudo R ²	0.063	0.066	0.069	0.067	0.067	0.066

Variable dependiente: ocupado (=1 si la mujer reporta empleo).

Restricciones de exclusión: tiene_hijos, hijos_3mas.

Errores estándar robustos con cluster por UPM. Datos deduplicados de ENOE.

*** p<0.01, ** p<0.05, * p<0.10.