



UNIVERSIDAD AUTÓNOMA DEL ESTADO DE MÉXICO

CENTRO UNIVERSITARIO UAEM VALLE DE MÉXICO

**Banca y seguros: brecha salarial de género, métrica
a través de la técnica Oaxaca Blinder**

TESIS

Que para obtener el Título de

LICENCIADA EN ACTUARÍA

P r e s e n t a

C. Mijangos Miranda Arlette Adamary

Asesora: Dra. en E. Jéssica Gámez Arroyo



Atizapán de Zaragoza, Edo. de Méx. octubre de 2026

ÍNDICE

RESUMEN	6
ABSTRACT	7
INTRODUCCIÓN.....	8
CAPÍTULO 1. MARCO TEÓRICO E HISTÓRICO SOBRE LOS SALARIOS Y LA BRECHA SALARIAL DE GÉNERO.....	11
1.1 Definición y medición de los salarios	11
1.2 Aportaciones desde la microeconomía	16
1.3 Teoría de capital humano	21
1.4 Factores que generan la diferenciación salarial.....	26
1.5 Perspectiva de género: brecha salarial y discriminación	29
1.6 Análisis histórico de los salarios y la brecha salarial de género.....	32
CAPÍTULO 2. ANÁLISIS DE REGRESIÓN CON DATOS DE CORTE TRANSVERSAL	39
2.1 Tipos de datos: series de tiempo, corte transversal y datos de panel.....	39
2.2 Modelo de regresión clásico	42
2.3 Variables binarias y categóricas	46
2.4 Modelos con interacciones entre variables cualitativas.....	50
2.5 Bondad de ajuste.....	53
2.6 Pruebas de diagnóstico	56
2.7 Técnica econométrica Oaxaca-Blinder.....	59
CAPÍTULO 3. CONSTRUCCIÓN E INTERPRETACIÓN DE LA BASE DE DATOS PARA MÉXICO Y PARA EL CASO DE ESTUDIO	64

3.1 México: análisis histórico de los salarios	64
3.2 Aportación de las actividades económicas de Banca y Seguros	69
3.3 Fuente de información: Encuesta Nacional de Ocupación y Empleo	76
3.4 Interpretación estadística y gráfica de variables de la Encuesta Nacional de Ocupación y Empleo, a nivel nacional	82
3.5 Análisis de la muestra seleccionada: Banca y Seguros, edad de 23 a 27 años.	91
CAPÍTULO 4. APLICACIÓN DE LA TÉCNICA ECONOMETRICA OAXACA-BLINDER PARA EL CASO DE ESTUDIO.	100
4.1 Selección de variables	100
4.2 Propuesta de modelo.....	102
4.3 Pruebas de diagnóstico	106
4.4 Resultados.....	119
4.5 Discusión: ventajas y limitaciones	123
CONCLUSIONES.....	126
ANEXOS.....	131
REFERENCIAS	144

ÍNDICE DE TABLAS

Tabla 1. Definiciones sobre el mercado laboral	18
Tabla 2. Distribución por sexo de la muestra depurada (ENOE 2025-I, nivel nacional).....	85
Tabla 3. Estadísticos descriptivos generales de la muestra depurada (ENOE 2025-I, nivel nacional)	86
Tabla 4. Salario mensual por sexo (ENOE 2025-I, nivel nacional)	87
Tabla 5. Nivel educativo por sexo (ENOE 2025-I, nivel nacional)	88
Tabla 6. Experiencia laboral potencial por sexo (ENOE 2025-I, nivel nacional)	90
Tabla 7. Composición por sexo (Banca y Seguros).....	92
Tabla 8. Estadísticos descriptivos de la muestra seleccionada: Banca y Seguros (23–27 años)	92
Tabla 9. Salario mensual por sexo (Banca y Seguros, 23–27 años).....	93
Tabla 10. Nivel educativo por sexo (Banca y Seguros, 23–27 años)	94
Tabla 11. Experiencia laboral potencial por sexo.....	97

ÍNDICE DE FIGURAS

Figura 1. Elementos que conforman el salario	13
Figura 2. Ejemplo de implementación en R	59
Figura 3. Distribución del nivel educativo por sexo (ENOE 2025-I, nivel nacional)	89
Figura 4. Distribución del nivel educativo por sexo en Banca y Seguros (23-27 años).....	96
Figura 5. Distribución de la muestra por ocupación y sexo (Banca y Seguros, 23-27 años).....	97
Figura 6. Resultados de la prueba Breusch-Pagan para hombres.....	107
Figura 7. Resultados de la prueba Breush-Pagan para mujeres.....	108
Figura 8. Resultados de la prueba Jarque-Bera para hombres.....	109

Figura 9. Resultados de la prueba Jarque-Bera para mujeres.....	110
Figura 10. Resultados de la prueba Factor de Inflación de Varianza (VIF) para hombres	112
Figura 11. Resultados de la prueba Factor de Inflación de Varianza (VIF) para mujeres .	113
Figura 12. Gráfico de residuos vs. valores ajustados (hombres).....	114
Figura 13. Histograma de residuos (hombres).....	115
Figura 14. Gráfico Q-Q de residuos (hombres).....	115
Figura 15. Gráfico de residuos vs. valores ajustados (mujeres)	116
Figura 16. Histograma de residuos (mujeres).....	117
Figura 17. Gráfico Q-Q de residuos (mujeres).....	117
Figura 18. Descomposición de la brecha salarial por componentes (modelo en logaritmos)	121
Figura 19. Coeficientes estimados por sexo (modelo en logaritmos).....	122

RESUMEN

La presente investigación analiza la brecha salarial de género en Banca y Seguros en México para la población de 23 a 27 años. El objetivo principal consiste en identificar si existen diferencias salariales entre hombres y mujeres y determinar qué proporción de dicha brecha puede explicarse por diferencias en capital humano y qué parte podría asociarse a factores estructurales o discriminatorios.

Para ello, se utilizaron datos de corte transversal provenientes de la Encuesta Nacional de Ocupación y Empleo (ENOE) del primer trimestre de 2025. La metodología empleada se basa en la descomposición Oaxaca-Blinder, complementada con modelos de regresión lineal y pruebas de diagnóstico econométrico para verificar la validez estadística del modelo.

Los resultados muestran la existencia de una brecha salarial significativa dentro del sector analizado. A pesar de que hombres y mujeres presentan niveles relativamente similares de escolaridad y experiencia laboral potencial, persisten diferencias importantes en los ingresos percibidos y en los retornos asociados al capital humano. La evidencia sugiere que una parte importante de la brecha salarial no se explica únicamente por diferencias en características observables, sino también por la manera en que el mercado laboral remunera dichas características entre hombres y mujeres.

Finalmente, la investigación aporta evidencia empírica sobre la persistencia de desigualdades salariales de género dentro de una actividad económica altamente profesionalizada, contribuyendo al análisis de la desigualdad laboral en México desde una perspectiva econométrica aplicada.

PALABRAS CLAVE: Brecha salarial de género, Oaxaca-Blinder, Capital humano, discriminación salarial, ENOE.

CLASIFICACIÓN: J16, J31, J71, G20

ABSTRACT

This study analyzes the gender pay gap in Mexico's banking and insurance sector for the population aged 23 to 27. The main objective is to determine whether wage differences exist between men and women and to identify what proportion of this gap can be explained by differences in human capital and what portion might be associated with structural or discriminatory factors.

To this end, cross-sectional data from the National Occupation and Employment Survey (ENOE) for the first quarter of 2025 were used. The methodology employed is based on the Oaxaca-Blinder decomposition, supplemented with linear regression models and econometric diagnostic tests to verify the statistical validity of the model.

The results show the existence of a significant wage gap within the sector analyzed. Although men and women have relatively similar levels of education and potential work experience, significant differences persist in earnings and in the returns associated with human capital. The evidence suggests that a significant portion of the wage gap is not explained solely by differences in observable characteristics, but also by the way in which the labor market rewards these characteristics among men and women.

Finally, the study provides empirical evidence on the persistence of gender wage gaps within a highly professionalized sector, contributing to the analysis of labor inequality in Mexico from an applied econometric perspective.

KEYWORDS: Gender pay gap, Oaxaca-Blinder, human capital, wage discrimination, ENOE.

INTRODUCCIÓN

La desigualdad salarial entre hombres y mujeres constituye uno de los problemas más persistentes y complejos dentro del mercado laboral contemporáneo. A pesar de los avances alcanzados en materia de educación, participación económica e inclusión laboral femenina, continúan existiendo diferencias significativas en los ingresos percibidos por hombres y mujeres, incluso cuando ambos presentan niveles educativos, puestos de trabajo y características laborales similares. Esta situación no solo refleja desigualdades económicas, sino también dinámicas estructurales asociadas con la discriminación, la segmentación ocupacional y los distintos retornos que el mercado laboral otorga a las características productivas de cada individuo.

La brecha salarial de género representa una problemática especialmente relevante en México, debido a las condiciones estructurales que caracterizan al mercado laboral nacional, tales como la informalidad, la desigual distribución de oportunidades laborales, la segregación ocupacional y las persistentes desigualdades de género. De acuerdo con el Instituto Mexicano para la Competitividad (IMCO, 2024), México ocupa la posición 109 de 146 países dentro del subíndice de Participación y Oportunidades Económicas del Índice Global de Brecha de Género, lo que evidencia que las desigualdades económicas entre hombres y mujeres continúan representando un reto significativo para el país.

Asimismo, distintos indicadores muestran que las mujeres continúan percibiendo menores ingresos que los hombres aun desempeñando actividades similares o contando con niveles comparables de escolaridad y experiencia laboral. Esta situación resulta particularmente preocupante debido a las implicaciones económicas y sociales que genera, ya que limita las oportunidades de crecimiento profesional, autonomía económica y acceso equitativo al desarrollo laboral para las mujeres.

Si bien la teoría del capital humano sostiene que las diferencias salariales pueden explicarse a partir de factores como la educación, la experiencia laboral y las habilidades adquiridas por

los individuos (Becker, 1964; Mincer, 1974), diversos estudios han demostrado que estas variables no son suficientes para explicar completamente la brecha salarial de género. En la práctica, las mujeres suelen enfrentar menores retornos asociados a su capital humano, aun cuando poseen características similares o incluso superiores a las de los hombres (Blau Y Kahn, 2017; Goldin, 2014). Esto sugiere la existencia de factores estructurales adicionales, relacionados con dinámicas de discriminación, roles de género y segmentación del mercado laboral.

En este contexto, la presente investigación tiene como objetivo analizar la brecha salarial entre hombres y mujeres en México, enfocándose específicamente en Banca y Seguros para la población de 23 a 27 años. La elección de este sector responde a su relevancia dentro del sistema financiero mexicano, así como a las características de profesionalización y especialización que presentan los empleos que lo conforman. Asimismo, el rango de edad seleccionado permite estudiar una etapa inicial de inserción laboral, en la cual comienzan a observarse diferencias salariales y trayectorias profesionales diferenciadas entre hombres y mujeres.

El análisis se desarrolla utilizando información proveniente de la Encuesta Nacional de Ocupación y Empleo (ENOE) correspondiente al primer trimestre de 2025, empleando datos de corte transversal para construir una muestra enfocada en trabajadores asalariados de Banca y Seguros. A partir de esta información, se consideran variables relacionadas con capital humano, experiencia laboral potencial y características sociodemográficas, con el propósito de identificar aquellos factores que influyen en la determinación de los ingresos laborales. Metodológicamente, la investigación utiliza la descomposición Oaxaca-Blinder (1973), técnica econométrica ampliamente utilizada en estudios de desigualdad salarial, la cual permite separar la brecha salarial observada en dos componentes: una parte explicada por diferencias en características observables —como escolaridad y experiencia laboral— y otra asociada a diferencias en los retornos que el mercado laboral otorga a dichas características. Esta metodología resulta particularmente útil para identificar qué proporción de la desigualdad salarial puede atribuirse a factores relacionados con capital humano y qué parte podría estar asociada a elementos estructurales o discriminatorios.

Los resultados obtenidos evidencian la existencia de una brecha salarial significativa dentro

del sector analizado. A pesar de que hombres y mujeres presentan niveles relativamente similares de escolaridad y experiencia laboral potencial, persisten diferencias importantes tanto en los ingresos percibidos como en los retornos asociados a estas variables. En particular, la evidencia sugiere que las mujeres reciben menores retornos a su capital humano, lo que indica que la brecha salarial no puede explicarse únicamente por diferencias en formación académica o experiencia laboral.

Finalmente, esta investigación busca contribuir al análisis de la desigualdad salarial de género en México desde una perspectiva econométrica y aplicada, aportando evidencia específica para Banca y Seguros. Asimismo, pretende generar elementos que permitan comprender con mayor profundidad los factores que intervienen en la formación de la brecha salarial y la persistencia de desigualdades dentro del mercado laboral mexicano, promoviendo la reflexión sobre la necesidad de construir entornos laborales más equitativos e incluyentes.

CAPÍTULO 1. MARCO TEÓRICO E HISTÓRICO SOBRE LOS SALARIOS Y LA BRECHA SALARIAL DE GÉNERO

La comprensión de la brecha salarial de género requiere partir de los principales enfoques teóricos que explican la determinación de los salarios y las desigualdades observadas en el mercado laboral. En este sentido, el presente capítulo reúne los fundamentos conceptuales que sustentan la investigación, abordando los conceptos de salario, las principales teorías salariales, la teoría del capital humano y los enfoques relacionados con la perspectiva de género y la discriminación salarial. Asimismo, se revisa la literatura especializada sobre brecha salarial de género con el propósito de construir el marco teórico que servirá de base para el análisis empírico desarrollado en los capítulos posteriores.

1.1 Definición y medición de los salarios

El salario o recompensa natural al trabajo, es la retribución que recibe un trabajador por la prestación de sus servicios en una actividad productiva (Smith, 1776). Se establece mediante acuerdos entre empleadores y empleados, influenciado por factores como la oferta y demanda de trabajo, la productividad laboral y las regulaciones gubernamentales (Borjas, 2016; Mankiw, 2014), este incluye, además del salario base, los complementos salariales asociados a la persona, al puesto, al desempeño, o a la realización de trabajo adicional, sumando tanto los importes en efectivo como las retribuciones en especie. Según el artículo 82 de la Ley Federal del Trabajo, el salario se define como "la retribución que debe pagar el patrón al trabajador por su trabajo" (Ley Federal del Trabajo, 1970, art. 82).

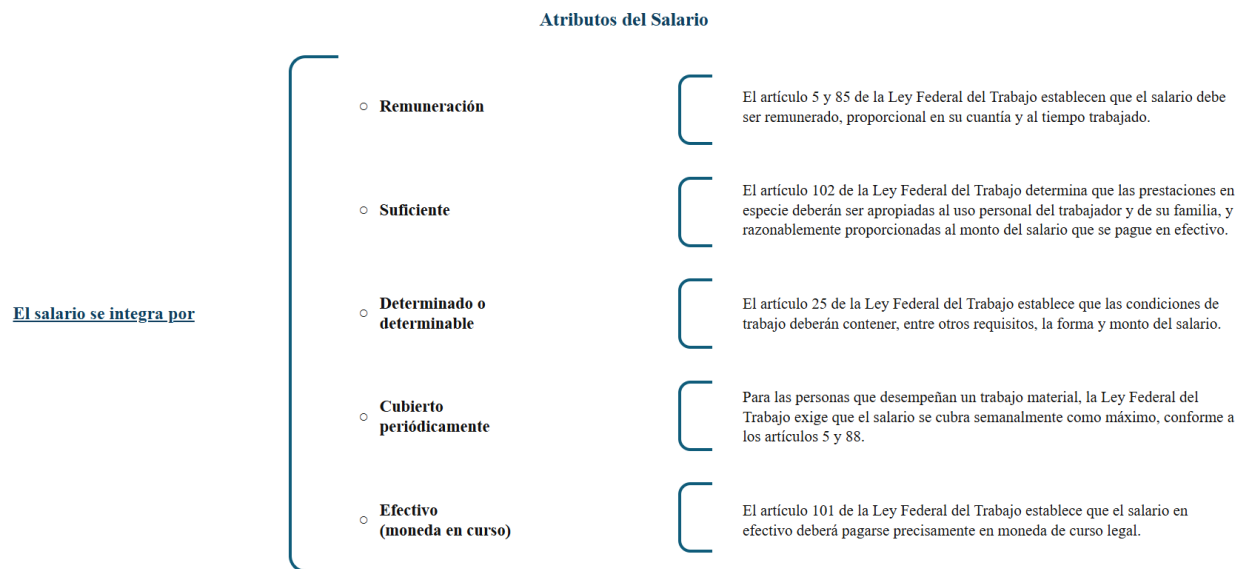
Es importante distinguir entre salario, sueldo y otras formas de remuneración. Aunque en el lenguaje cotidiano suelen utilizarse como sinónimos, el salario generalmente se asocia con una retribución calculada en función del tiempo trabajado, la producción realizada o la prestación efectiva del servicio, mientras que el sueldo suele referirse a una cantidad fija pagada de manera periódica, comúnmente de forma mensual, independientemente de variaciones en la producción. Asimismo, existen otras modalidades de remuneración

reconocidas en el ámbito laboral, como el salario por unidad de tiempo, por unidad de obra o a destajo, por comisión y por tarea, las cuales permiten adaptar la forma de pago a las características específicas de cada actividad económica (Ley Federal del Trabajo, 1970).

Existen diversas formas de determinar y pagar el salario. Una de ellas es el salario por unidad de tiempo, en el cual la remuneración se establece con base en horas, días o meses trabajados. También existe el salario por unidad de obra o a destajo, donde el pago depende de la cantidad de trabajo realizado o de las unidades producidas. Asimismo, pueden existir esquemas por comisión, frecuentes en actividades comerciales y de ventas, donde la remuneración está vinculada a los resultados obtenidos. Por otra parte, el salario puede clasificarse como ordinario o extraordinario; el primero corresponde a la retribución derivada de la jornada normal de trabajo, mientras que el segundo se genera por conceptos adicionales, como horas extraordinarias o labores realizadas en días de descanso.

Finalmente, se distingue entre salario nominal y salario real. El salario nominal corresponde a la cantidad monetaria que recibe el trabajador por la prestación de sus servicios, mientras que el salario real representa el poder adquisitivo de dicho ingreso, es decir, la cantidad de bienes y servicios que pueden adquirirse con él una vez considerado el efecto de los precios y la inflación (Mankiw, 2014).

Figura 1. Elementos que conforman el salario



Fuente. Elaboración propia, con información de Moreno (Moreno, 2018).

La medición de salarios hacia las mujeres enfrenta un desafío importante en cuanto a discriminación salarial, ya que históricamente han ganado menos que los hombres por realizar tareas similares. Esta brecha salarial de género ha sido un tema central en la discusión sobre igualdad de derechos laborales en México y en otros países. Se conoce que la brecha salarial de género es el porcentaje resultante de dividir dos cantidades: la diferencia entre el salario de los hombres y las mujeres, dividida entre el salario de los hombres (ONU MUJERES, 2025).

Las cantidades que forman parte del salario deben incluir los impuestos y las aportaciones de la persona trabajadora al sistema de seguridad social y pensiones. El salario puede clasificarse en diversas categorías, tales como salario nominal que es la suma de dinero en la moneda pesos de curso legal que recibe periódicamente un trabajador por parte de su empleador por lo que puede fijarse por unidad de tiempo, tareas realizadas, por comisión o cualquier otra forma acordada por las partes. Mientras que el salario real se obtiene al descontar el efecto de la inflación y se calcula dividiendo el salario nominal entre el índice

nacional de precios al consumidor (Krugman Y Wells, 2018). Una de sus categorías es el salario mínimo, entendido como la remuneración mínima que, por disposición legal, debe recibir una persona trabajadora por los servicios prestados durante su jornada laboral. Su monto es determinado y actualizado por la Comisión Nacional de los Salarios Mínimos (CONASAMI) conforme a las condiciones económicas y sociales del país (CONASAMI, 2025). Igualmente, otras de sus clasificaciones es el salario por productividad, salario por hora, semanal o mensual, salario monetario y no monetario. Se conocen otras retribuciones las cuales no las cuentan como salario, pero son cantidades percibidas por un trabajador las cuales son: indemnizaciones o suplidos por los gastos realizados de su actividad laboral, prestaciones e indemnizaciones de la Seguridad Social, el plus de distancia y transporte, indemnizaciones correspondientes a traslados, indemnizaciones correspondientes a suspensiones o despidos y por último por alojamiento y manutención (Moreno, 2018).

En la presente investigación se abordan tres corrientes teóricas que explican la formación de los salarios: la teoría clásica (no admite brechas salariales), la teoría del capital humano (admite brechas salariales por formación de capital humano), y las ideologías basadas en perspectiva de género y discriminación (admite brechas salariales por cuestiones de género).

Neffa (2007) explica que, según la teoría económica, el salario refleja el precio de la mano de obra, el cual difiere de su costo por unidad producida, en el sentido que este último toma en cuenta la eficiencia con que se utiliza el trabajo. Se destacan por lo tanto dos conceptos; por un lado, el salario, que es la cantidad nominal que se paga al trabajador, y, por otro lado, el costo por unidad, en el cual también se considera la productividad, es decir, la eficiencia del trabajo.

La teoría clásica de los salarios, desarrollada por Adam Smith, sostiene que los salarios se determinan mediante la interacción entre la oferta y la demanda en el mercado laboral (Smith, 1776). En este mercado, los trabajadores representan la oferta de trabajo, mientras que los empleadores o empresas constituyen la demanda de trabajo. Desde esta perspectiva, cuando la demanda de trabajadores supera la oferta disponible, los salarios tienden a incrementarse;

por el contrario, cuando la oferta de trabajo es mayor que la demanda, los salarios tienden a disminuir. Esta visión considera que tanto los trabajadores como los empleadores buscan maximizar sus beneficios y que el mercado laboral tiende a ajustarse de manera natural hasta alcanzar un equilibrio entre la oferta y la demanda de mano de obra. En consecuencia, este enfoque no contempla la existencia de una brecha salarial, ya que supone que cualquier diferencia en las remuneraciones se corrige mediante el funcionamiento del propio mercado.

En contraste, la teoría del capital humano sostiene que la educación, la formación y las habilidades humanas del individuo son factores que determinan su salario (Becker, 1964; Mincer, 1974; Schultz, 1961). Según esta perspectiva, aquellos que invierten en su educación y formación desarrollan habilidades que los hacen más productivos en el ámbito laboral, lo cual les permite acceder a mejores salarios. Economistas como Theodore Schultz, Gary Becker y Mincer fueron los principales autores en desarrollar esta teoría. En esta corriente teórica da cabida a las brechas salariales, posibles cuando existen diferencias en la formación de cada individuo (Gámez, Rosas Y Lapa, 2022); es decir, si un hombre obtiene un mayor ingreso que una mujer, es porque el hombre ha invertido más en su formación profesional. Sin embargo, esta teoría ha sido criticada, ya que en la práctica no siempre se ajusta a la realidad de las desigualdades salariales entre géneros, lo cual sugiere que otros factores, como la discriminación, podrían estar influyendo en las disparidades salariales.

Por otro lado, las ideologías centradas en el género y la discriminación ofrecen un enfoque que integra dimensiones sociales, políticas y jurídicas para visibilizar y combatir las desigualdades que enfrentan las mujeres en el ámbito laboral y en otros espacios de la sociedad. Asimismo, buscan promover una sociedad basada en la igualdad de derechos y oportunidades para todas las personas. La economista Corina Rodríguez argumenta que dentro de este marco de pensamiento se desarrolla el concepto de economía feminista, el cual analiza la participación económica de las mujeres y pone de manifiesto los mecanismos de discriminación presentes en el mercado laboral, permitiendo comprender cómo las estructuras y prácticas económicas contribuyen a la persistencia de las desigualdades de género.

Las teorías expuestas ofrecen distintas perspectivas para comprender la determinación de los salarios y las causas de la brecha salarial entre hombres y mujeres. Mientras que la teoría clásica atribuye la formación de los salarios al equilibrio entre la oferta y la demanda de trabajo, la teoría del capital humano sostiene que las diferencias salariales responden a las inversiones realizadas en educación, experiencia y capacitación. Por su parte, la perspectiva de género y discriminación plantea que, aun cuando hombres y mujeres poseen características productivas similares, pueden persistir diferencias en las remuneraciones derivadas de factores estructurales, sociales e institucionales. En este sentido, la presente investigación adopta principalmente el enfoque de la teoría del capital humano como base para la especificación econométrica, al considerar variables como la escolaridad, la experiencia laboral potencial y las horas trabajadas. No obstante, también incorpora la perspectiva de género para reconocer que una parte de la brecha salarial puede no ser explicada únicamente por las diferencias en capital humano, sino por factores asociados a la discriminación o a retornos diferenciados de dichas características. Esta integración resulta congruente con la metodología de Oaxaca-Blinder, la cual permite distinguir entre la porción de la brecha explicada por diferencias observables y aquella que permanece sin explicar.

Neffa (2007) explica que los salarios fijados por el mercado y ofrecidos por los empleadores a nivel sectorial, se corresponden con la información disponible sobre el mercado de trabajo y con lo que consideran su certera estimación *ex ante* de la productividad marginal del trabajo, aportada por el último trabajador incorporado. Dicho argumento resulta relevante, ya que, en esta investigación se considera el caso de estudio del sector bancario y de seguros, donde los niveles salariales pueden ser resultado de los conocimientos y habilidades requeridos en estas actividades económicas.

1.2 Aportaciones desde la microeconomía

Actualmente la economía se subdivide en dos áreas fundamentales: la microeconomía y la macroeconomía. Generalmente se considera a Adam Smith como el padre de la economía

moderna y fundador de la microeconomía, rama de la economía que hoy se ocupa del comportamiento de entidades individuales tales como mercados, empresas y hogares (Samuelson y Nordhaus, 2010), la microeconomía es ocupada para los conductos de unidades económicas individuales (Varian, 2014). Estas unidades son los consumidores, los trabajadores, los inversores, los propietarios de tierra, las empresas en realidad, cualquier individuo o entidad que desempeñe algún papel en el funcionamiento de nuestra economía. Es por ello por lo que la microeconomía explica cómo y por qué estas unidades toman decisiones económicas.

En microeconomía los salarios representan el precio del trabajo (Borjas, 2016). Por lo que el análisis microeconómico de los salarios se centra en cómo se determinan, cómo influyen las políticas laborales y cómo se distribuyen los ingresos entre los individuos. Entre los conceptos más importantes destacan cuatro puntos el primero es la oferta y demanda de trabajo el cual se divide en dos, demanda de trabajo y oferta de trabajo; se entiende por, demanda de trabajo a la cantidad de trabajo que las empresas están dispuestas a ofrecer diferentes niveles salariales (Borjas, 2016; Mankiw, 2014). Generalmente, si el salario aumenta, la demanda de trabajo disminuye, ya que las empresas podrían contratar menos trabajadores a un precio más alto. Mientras que la oferta, es la cantidad de trabajo que los individuos están dispuestos a ofrecer a diferentes niveles salariales (Borjas, 2016; Mankiw, 2014). Si el salario mejora, la oferta de trabajo tiende a aumentar porque más personas estarán dispuestas a trabajar.

Como segundo punto, se encuentra el equilibrio en el mercado laboral por lo que el salario de equilibrio consiste en el salario al cual la cantidad de trabajo que las empresas desean contratar es igual a la cantidad de trabajo que los trabajadores están dispuestos a ofrecer. Si hay un desequilibrio (por ejemplo, salario mínimo o salario por debajo del equilibrio), pueden surgir desempleo o escasez de trabajadores (Mankiw, 2014).

Resulta relevante incorporar los factores que contribuyen a la formación de los salarios; se compone de tres elementos fundamentales: habilidades y educación, los salarios tienden a

ser más altos en trabajos para quienes poseen grados académicos superiores y/o habilidades superiores. Esto se conoce como la teoría de capital humano, en la cual destaca que las personas son consideradas inversionistas, es decir, la formación se considera una inversión, cuyo retorno se obtiene en el futuro (Becker, 1964; Mincer, 1974).

Uno de los elementos fundamentales es el concepto de Productividad Marginal del Trabajo (PML), ya que en la formación del salario se puede considerar el valor que un trabajador adicional aporta a la producción total de la empresa, por lo tanto, entre mayor sea la productividad marginal, puede haber mejoras en el ingreso del individuo (Varian, 2014). Finalmente se encuentra la Brecha Salarial en esta se puede limitar las oportunidades de ascensos y desarrollo profesional para ciertos grupos, lo que a su vez se afecta de manera negativa a sus salarios (Blau Y Kahn, 2017).

Las aportaciones en microeconomía suelen relacionar con las contribuciones de los trabajadores en la economía. Por medio de su trabajo y la producción de bienes y servicios. Por lo tanto, la microeconomía principalmente se enfoca en el estudio de las decisiones económicas como los consumidores y las empresas. En consecuencia, el salario es relevante de estudiar, ya que es el precio que se paga por laborar. Entre las principales aportaciones identificadas destacan las que se presentan en la Tabla 1.

Tabla 1. Definiciones sobre el mercado laboral

APORTACIÓN	DEFINICIÓN
Oferta y demanda de trabajo	Determina el nivel salarial a partir de la interacción entre empleadores y trabajadores en el mercado laboral.
Equilibrio laboral	Punto en el que la oferta y demanda de trabajo se igualan, estableciendo un salario de mercado.
Capital humano	Conjunto de habilidades, educación y experiencia que incrementan la productividad de los individuos.
Productividad marginal	Valor adicional que aporta un trabajador a la producción total, utilizado para determinar su remuneración.
Brecha salarial	Diferencia en ingresos entre grupos poblacionales que no puede explicarse únicamente por características productivas observables.

Fuente. Elaboración propia con información de Samuelson y Nordhaus (2010)

Un tema importante de la microeconomía es el papel de los precios. Todas las disyuntivas antes descritas se basan en los precios a los que se enfrentan los consumidores, los trabajadores o las empresas (Varian, 2014). Por ejemplo, un consumidor intercambia carne de pescado por carne de res, calculando, en parte, en sus preferencias por cada uno, pero también en sus precios. Asimismo, los trabajadores intercambian trabajo, en parte, el «precio» que pueden percibir por su trabajo, es decir, en el salario. Mientras que las empresas deciden contratar o no, más trabajadores o comprar más máquinas calculando, en parte, en los salarios y en los precios de las máquinas.

La microeconomía permite analizar el comportamiento económico de algunas empresas y da un punto de vista en la toma de decisiones, tanto a largo plazo como a corto plazo (Pindyck Y Rubinfeld, 2018). No obstante, es importante destacar que los modelos de la teoría microeconómica son representaciones simplificadas de la realidad. Utilizando herramientas como la teoría de la demanda y oferta, podemos predecir qué pasara con los precios de los bienes y servicios, que necesitemos o tengamos, ante un cambio en la tasa impositiva.

La microeconomía apoya en la toma de decisiones, pues permite a los agentes comprender mejor el comportamiento económico del mercado y analizar cómo responden los consumidores ante el cambio de precios o la cantidad de oferentes (Pindyck Y Rubinfeld, 2018). Así que la microeconomía da una estructura básica de análisis y sus datos brindan un marco esencial para la toma de decisiones en diferentes áreas de una empresa como *marketing*, diseño, ventas o aplicación de estrategias. Por lo tanto, repercute positivamente en el crecimiento económico de la misma. También la microeconomía analiza las interacciones entre los diferentes mercados, lo que permite predecir qué pasará si determinadas variables se modifican. Esto es de utilidad para organizar comportamientos, decisiones y preferencias y hacer frente a los cambios de otros agentes económicos. Con esto queda claro que es una de las herramientas más importantes y útiles para poder comprender el funcionamiento económico de la sociedad.

Neffa (2007) realiza una crítica a la teoría neoclásica ortodoxa, pues asegura que desconoce el hecho de que, del lado de la oferta de trabajo, el mercado laboral se comporte de una manera diferente al paradigma competitivo, por diversas razones:

- Dentro de una misma región, y sobre todo entre regiones, existe normalmente una gran variedad de salarios –incluso para tareas similares–, pero la movilidad de los trabajadores hacia los empleos más remunerados no es automática, pues significa asumir un riesgo y tener que enfrentar serias restricciones, no sólo de tipo económico (tanto Adam Smith como Marx habían justificado la existencia de diferencias salariales en función de la desutilidad del trabajo realizado en condiciones penosas o insalubres y del costo de reproducción de la fuerza de trabajo para diversas profesiones o en diversas situaciones).
- En el nivel agregado de la rama de actividad la oferta de trabajo aumentará sólo si aumentan los salarios ofrecidos, pero en el nivel de la empresa individual, sobre todo si es de pequeña dimensión, la oferta de trabajo tiene una posición casi horizontal pues involucra a pocos trabajadores. Entonces, cuando se requiere incorporar a pocos trabajadores nuevos, la empresa no necesitará incrementar los salarios que fija el mercado para reclutar más mano de obra, siempre que sus calificaciones sean similares.
- Los trabajadores no conocen todas las posibilidades en materia de empleos y de salarios disponibles. Su información es más bien escasa y no procuran actualizarla en permanencia, por lo cual no pueden desplazarse hacia los empleos mejor remunerados.
- Esta teoría no considera explícitamente el rol de la familia cuando ésta adopta la decisión de oferta de trabajo de sus miembros, en el sentido de elegir frente a otros usos alternativos no remunerados –las tareas domésticas, por ejemplo–.
- Si se lleva el razonamiento hasta su límite, y simplificando, no existe un diagrama único de oferta de trabajo. Se debe tener en cuenta la segmentación de los mercados y prestar atención a lo que sucede en cada unidad productiva. Pero entre todos esos

niveles existe una interdependencia real: si las diferencias salariales entre ramas o empresas son consideradas excesivas por los trabajadores de una rama o empresa mal paga, los empleadores que pagan poco deben reducir tales diferenciales si desean retener la mano de obra.

Estos argumentos resultan oportunos en la presente investigación, ya que se resalta el hecho de que los salarios en la práctica no son estimados únicamente como reflejo de la productividad laboral, así como las brechas salariales entre regiones y entre sectores; también da cuenta de las restricciones que pueden enfrentar las mujeres en el rol familiar. En este sentido Neffa (2007) concluye:

Además de la situación propia del mercado de trabajo, intervienen otros factores diversos en la determinación del nivel salarial, en especial: la estructura del sistema productivo nacional y su inserción en la división internacional del trabajo, el ritmo y el modelo de crecimiento económico, la correlación de fuerzas existente entre los interlocutores sociales, las políticas públicas macroeconómicas en materia de remuneraciones, salario indirecto, distribución del ingreso, así como el nivel de los excedentes económicos, los modos de gestión de la fuerza de trabajo y las estrategias y estructura de la organización sindical (Neffa, 2007, p. 88).

1.3 Teoría de capital humano

Las capacidades humanas son consideradas como un tipo de capital, en la tradición neoclásica ya se encontraba en Alfred Marshall, quien estudia la influencia de la demanda en la distribución del ingreso (Marshall, 2013). En ese tiempo existía un elemento importante para la explicación de las diferencias salariales, según Marshall es la existencia de una relación complementaria entre el trabajo y el capital; de este modo, el incremento de capital no solo en la cantidad, sino en la complejidad tecnológica, implicando un incremento en la cantidad y calidad de trabajo demandado (Marshall, 2013)

Debido de lo previamente mencionado, Marshall denomina salario de eficiencia a la

remuneración por unidad de eficiencia del trabajo; es decir, los salarios medidos no como la remuneración de trabajo por tarea con referencia a la cantidad de producción resultante, sino con referencia al esfuerzo de habilidad y eficiencia que obtiene un trabajador; por lo tanto, salarios de eficiencia de Marshall muestra que provocan que paguen diversos salarios a los trabajadores con diferente nivel de eficiencia. En consecuencia, se introduce la competencia por el nivel de educación de los individuos como punto crítico de la diferencia en salarios de los trabajadores (Marshall, 2013).

Durante la primera mitad del siglo XX, diversos autores retomaron la discusión sobre las diferencias salariales entre tipos de trabajo, incorporando conceptos como educación, capacitación y tasa de retorno de la inversión en capital humano (Mincer, 1958; Schultz, 1961; Becker, 1964).

La teoría del capital humano sostiene que la educación, la formación y las habilidades del individuo constituyen factores determinantes de su productividad y, por tanto, de su salario (Becker, 1964; Mincer, 1974; Schultz, 1961). Según esta perspectiva, aquellos que invierten en su educación, formación y experiencia laboral, desarrollan habilidades que los hacen más productivos en el ámbito de trabajo, lo cual les permite acceder a mejores salarios, de esta forma se entiende que indiscutiblemente los trabajadores que realmente deseen mejorar su posición económica dentro del sistema actuarán de una manera racional al momento de elegir invertir en más educación y formación profesional. En términos generales, la teoría del capital humano sostiene que una mayor inversión en educación, capacitación y formación incrementa la productividad de los individuos y, en consecuencia, sus posibilidades de obtener mejores remuneraciones. Bajo este enfoque, Gary Becker argumentó que la inversión en capital humano es comparable a la inversión en capital físico, debido a que ambas generan rendimientos económicos futuros (Becker, 1964).

Theodore Schultz influyó significativamente en el desarrollo de la teoría del capital humano

al considerar que el gasto destinado a la educación constituye una inversión y no un consumo. Desde esta perspectiva, la educación genera un valor agregado para la economía y permite que los individuos incrementen su productividad, lo que se traduce en mayores ingresos a lo largo de su vida laboral (Schultz, 1961). Por su parte, Jacob Mincer fue pionero en el análisis de la relación entre el capital humano y la distribución de los ingresos. Su modelo plantea que la educación y la capacitación en el trabajo incrementan la productividad de los individuos y, en consecuencia, sus salarios. Asimismo, analiza cómo las diferencias en la inversión en capital humano pueden contribuir a explicar la desigualdad observada en las remuneraciones entre los trabajadores (Mincer, 1974).

El capital humano puede entenderse como un activo intangible integrado por conocimientos, habilidades, salud y experiencia, elementos que inciden en la productividad individual y en el crecimiento económico (Becker, 1964; Schultz, 1961). Desde esta corriente teórica, las diferencias salariales pueden explicarse por los distintos niveles de inversión que cada individuo realiza en su formación y desarrollo de capital humano. En este sentido, si una persona percibe un mayor ingreso que otra, ello respondería, desde esta perspectiva, a que ha invertido más en educación, capacitación o experiencia laboral (Gámez, Rosas Y Lapa, 2022).

Sin embargo, esta teoría ha sido criticada, ya que en la práctica no siempre se ajusta a la realidad de las desigualdades salariales entre géneros, lo cual sugiere que otros factores, como la discriminación, la educación formal, capacitación en el trabajo, salud, movilidad laboral, tecnología y digitalización, podrían estar influyendo en las disparidades salariales (Blau Y Kahn, 2017). Esto mejoraría su rendimiento laboral y la competencia dentro del mercado de trabajo garantizaría una recompensa equivalente a su mayor capacitación, de este modo se deja todo lo relacionado a la inversión en capital humano en manos del trabajador.

Con lo ya mencionado el Capital Humano se considera un activo intangible, pero cuenta con un valor importante en el crecimiento económico. Existen elementos que son fundamentales para esta teoría como lo es inversión en educación y formación asignar en recursos

financieros y humanos para mejorar las habilidades conocimientos y competencias a las personas con el objetivo de poder aumentar su productividad y capacidad para contribuir no solo en la mejora las habilidades de las personas, sino que también aumenta su capacidad para generar ingresos y contribuye a la competitividad de una economía.

El desarrollo de habilidades y competencias es un elemento fundamental ya que es el desarrollo constante de habilidades profesionales fundamental para mantenerse competitivo y eficiente en cualquier laboral junto con técnicas que los trabajadores adquieren a lo largo de su vida laboral. La salud también forma parte del capital humano, ya que una población saludable puede mantener una mayor productividad, prolongar su vida laboral y contribuir de manera más eficiente al crecimiento económico (Schultz, 1961; Banco Mundial, 2020). Los retornos económicos son aquellos que invierten en su educación y habilidades tienen acceso a mejores trabajos, percibir salarios más altos y contar con una mayor estabilidad financiera. Si hablamos a nivel macroeconómico, la acumulación de capital humano en la población de México puede impulsar el crecimiento económico, aumentar la productividad, generar mayores ingresos nacionales y contribuir al desarrollo económico del país.

La teoría del capital humano también se relaciona con la movilidad social, ya que la educación y el desarrollo de habilidades pueden ampliar las oportunidades económicas y laborales de los individuos (OCDE, 2018). El cambio puede implicar mejoras o retrocesos en riqueza, educación, ocupación, ingresos, salud, entre otras dimensiones sociales y económicas. La movilidad social refleja el nivel de igualdad de oportunidades entre la población a mayor igualdad de oportunidades, mayor movilidad social.

El impacto de la Teoría de Capital Humano en la formación profesional superior es más visible en países con mejores niveles de vida, en aquellos en los que hay más ingreso per cápita y en donde el nivel de formación profesional es superior. Organismos de alcance mundial coinciden en enfatizar la importancia del Capital Humano, tales como la Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura (UNESCO), quien enfatiza la trascendencia del conocimiento y la formación de Capital Humano como

los fundamentos necesarios para el crecimiento de las naciones.

Para obtener un resultado de lo anterior, es indispensable fortalecer los sistemas de investigación e innovación, la creación de capacidad, el uso de tecnologías y las redes de científicos, la colaboración entre la universidad y el sector industrial (Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura - UNESCO, 2020). El Banco Mundial, propuso medir la posición de los países con base en una serie de indicadores que están integrados en cuatro dimensiones: 1. funcionamiento del régimen de incentivos económicos, 2. nivel de desarrollo de la educación y el Capital Humano, 3. capacidades de innovación y 4. uso de Tecnologías de la Información. Ese organismo refiere que México junto a Chile, Colombia, España, Perú y Portugal, han mostrado una trayectoria de progreso, Sin embargo, el avance hacia la economía global es muy reducido debido a diversos problemas a los cuales deben enfrentarse todos los países (Banco Mundial, 2020).

Por lo que existen diversas situaciones que influyen en el freno del desarrollo de las naciones subdesarrolladas, sin embargo, quien cuenta los resultados desfavorables, es el asociado a la inversión que se realiza en la educación superior. Domínguez señaló que los beneficios del avance científico y tecnológico se distribuyen de manera desigual entre países y grupos sociales, debido a las diferencias en recursos destinados a educación, ciencia e innovación (Domínguez, 2015).

Con ello reconociendo la simbiosis del Capital Humano y las instituciones educativas para el desarrollo de las naciones, Brunner y Ganga sostienen que las instituciones de educación superior desempeñan un papel fundamental en el desarrollo del capital humano, la competitividad y la superación de la pobreza (Brunner Y Ganga, 2016). Muestran la participación efectiva de Latinoamérica en la geopolítica del conocimiento avanzado, así como del gasto público y privado en la investigación y desarrollo los cuales son débiles. Este fenómeno ocurre casi en cualquiera dimensión, en la que intervenga el Producto Interno Bruto.

En un contexto de restricciones económicas y sociales, la teoría del capital humano sostiene que la educación constituye una inversión que incrementa las capacidades productivas de los individuos y amplía sus oportunidades de inserción en el mercado laboral. En este sentido, el acceso a la educación superior favorece la acumulación de capital humano y contribuye a mejorar las condiciones socioeconómicas de las personas. Por ello, resulta fundamental promover la igualdad de oportunidades educativas y fortalecer las estrategias que faciliten el acceso y la permanencia de los estudiantes en este nivel de formación (Martí-Noguera, Licandro, y Gaete-Quezada, 2018).

1.4 Factores que generan la diferenciación salarial

La diferenciación salarial es uno de los fenómenos más complejos dentro del ámbito laboral, ya que responde a una multiplicidad de factores económicos, sociales y estructurales que interactúan entre sí para determinar las condiciones de remuneración de las personas trabajadoras (Blau Y Kahn, 2017). Entender este fenómeno implica analizar las variables que inciden en que dos personas puedan recibir salarios distintos pese a desarrollar actividades similares o contar con un perfil profesional comparable. A continuación, se exponen los principales factores que inciden en la diferenciación salarial, respaldados por investigaciones recientes y análisis académicos.

Uno de los elementos fundamentales que explican las diferencias salariales es el capital humano. Este concepto, ampliamente desarrollado por la teoría económica clásica y neoclásica, hace referencia al conjunto de habilidades, conocimientos, formación académica y experiencia profesional que una persona posee.

De acuerdo con la teoría del capital humano, la inversión en educación y experiencia laboral se traduce en mayores niveles de productividad y, por ende, en mejores ingresos (Becker, 1964; Mincer, 1974). Por ejemplo, la diferencia de ingresos entre un administrativo y un corredor de bolsa puede explicarse por la complejidad y riesgo del trabajo, la formación requerida y la responsabilidad inherente al puesto. De esta manera, el capital humano continúa siendo una variable clave para entender las desigualdades salariales legítimas.

No obstante, además del nivel de formación, las condiciones laborales específicas también influyen significativamente en la retribución económica. Aquellos empleos que se desarrollan en entornos riesgosos, con horarios poco convencionales o que demandan una carga física o emocional elevada, suelen estar mejor remunerados como una forma de compensación. Asimismo, los complementos salariales, tales como pagos por productividad, bonos por antigüedad o beneficios asociados a ciertos logros, también generan distinciones en la estructura de ingresos de los trabajadores (OCDE, 2018). Estos factores hacen que incluso dentro de una misma empresa o sector, existan diferencias salariales entre empleados con puestos similares.

Sin embargo, no todas las disparidades pueden justificarse por causas objetivas o técnicas. Uno de los aspectos más problemáticos y persistentes en la estructura salarial de muchas sociedades es la discriminación laboral, especialmente la basada en el género. Diversos estudios han señalado que, incluso con igual formación y experiencia, las mujeres tienden a percibir ingresos menores que sus pares hombres. Esta brecha no puede explicarse únicamente a través del capital humano, sino que responde a sesgos culturales, estereotipos y estructuras laborales que privilegian a ciertos grupos sobre otros. Diversos estudios han documentado que estos sesgos se manifiestan tanto en los procesos de contratación como en las decisiones de promoción y asignación salarial (Blau Y Kahn, 2017).

En esta misma línea, la segregación ocupacional también contribuye a perpetuar las diferencias salariales. Este concepto se refiere a la concentración de ciertos grupos, especialmente mujeres, en sectores o actividades económicas que, históricamente, han sido menos valoradas y, por ende, peor remuneradas. Sectores como el educativo, el cuidado de personas mayores o la atención infantil, donde la presencia femenina es mayoritaria, suelen ofrecer salarios más bajos que áreas dominadas por hombres, como la ingeniería o la tecnología. Esta distribución desigual de la fuerza laboral no solo refleja patrones socioculturales, sino también decisiones políticas y económicas que configuran los niveles de remuneración en función del tipo de trabajo y no necesariamente de su valor social o económico (Anker, 1997; Blau Y Kahn, 2017).

Otro factor relevante es el impacto de la maternidad y las responsabilidades familiares sobre la trayectoria profesional de las mujeres. A diferencia de sus colegas varones, las mujeres enfrentan mayores obstáculos para mantener una continuidad laboral, debido a que las responsabilidades de cuidado recaen desproporcionadamente sobre ellas. Como consecuencia, muchas mujeres deciden incorporarse a ocupaciones de menor remuneración (o se ven forzadas a hacerlo) en empleos a tiempo parcial, licencias prolongadas o incluso interrupciones en su carrera. Estas decisiones, aunque necesarias desde el punto de vista familiar, repercuten en la progresión salarial, la acumulación de antigüedad y las oportunidades de ascenso, lo que genera una brecha que se amplía con el paso del tiempo (Goldin, 2014)

La falta de transparencia en las políticas salariales dentro de las organizaciones constituye otro obstáculo importante. En muchos entornos laborales no existen criterios claros ni públicos para la asignación de salarios, lo que permite la reproducción de desigualdades sin que estas sean detectadas o cuestionadas. La opacidad salarial impide a las personas trabajadoras comparar sus condiciones con las de sus colegas y dificulta la identificación de posibles prácticas discriminatorias. Estudios recientes indican que la transparencia salarial es una herramienta eficaz para reducir la brecha de género y promover entornos más equitativos (OCDE, 2021).

Además, los roles de género y los estereotipos sociales continúan influyendo en las decisiones profesionales de hombres y mujeres, así como en la valoración de los trabajos. A menudo, los empleos realizados mayoritariamente por mujeres son percibidos como extensiones de las tareas domésticas, y por tanto, subvalorados en términos económicos. Esta percepción cultural incide directamente en los niveles salariales asignados a dichos trabajos. Por otro lado, los hombres tienden a ocupar una mayor proporción de puestos de liderazgo o alta dirección, lo cual incrementa la brecha en los niveles más altos de remuneración (Blau Y Kahn, 2017).

El tipo de contrato y la jornada laboral también tienen un impacto directo sobre los ingresos. Las mujeres tienen una presencia significativamente mayor en empleos temporales, informales o de medio tiempo, lo que las deja en condiciones de mayor vulnerabilidad laboral

y limita su acceso a beneficios complementarios como seguridad social, vacaciones pagadas o bonos de desempeño. Estas modalidades contractuales, más precarias y menos estables, impiden la acumulación de derechos laborales y afectan de forma directa el nivel de ingreso promedio de las mujeres en comparación con los hombres (ILO, 2018).

Finalmente, la carencia de políticas públicas y organizacionales que promuevan la igualdad de género en el ámbito laboral perpetúa estas desigualdades. Las empresas e instituciones que no cuentan con medidas para garantizar la equidad salarial, la conciliación entre vida personal y profesional, y la eliminación de barreras estructurales, contribuyen a la reproducción de un modelo laboral injusto. La implementación de políticas de igualdad no solo es una cuestión de justicia social, sino también una estrategia que mejora el clima organizacional aumenta la productividad y fomenta un crecimiento económico más inclusivo (Banco Mundial, 2020).

La diferenciación salarial responde a un entramado complejo de factores objetivos, como el capital humano y las condiciones laborales, pero también a elementos subjetivos y estructurales como la discriminación, la segregación ocupacional y la falta de transparencia. Abordar esta problemática requiere un enfoque integral que combine políticas públicas efectivas, reformas organizacionales y un cambio cultural profundo que reconozca el valor del trabajo en función de su impacto real y no de los estereotipos asociados a quien lo realiza.

En conclusión, la diferenciación salarial en México es un fenómeno complejo influenciado por una variedad de factores económicos, sociales y culturales. A pesar de los esfuerzos realizados en las últimas décadas, la brecha salarial de género persiste como una de las manifestaciones más evidentes de esta desigualdad.

1.5 Perspectiva de género: brecha salarial y discriminación

Esta sección inicia con el concepto de igualdad de oportunidades, el cual se fundamenta en la firme convicción de que la genuina libertad se garantiza mediante la eliminación de las barreras de discriminación y del prejuicio (Gimeno, 1994, como se citó en Cadena, 2020). Comprender este concepto resulta un pilar en el tema de brechas salariales, ya que las mujeres

y los hombres deben competir en las mismas condiciones en el mercado laboral.

La brecha salarial de género se define como la diferencia relativa entre los ingresos promedio de hombres y mujeres, generalmente expresada como el porcentaje en que los ingresos de las mujeres son menores respecto a los de los hombres (OCDE, 2018; Blau y Kahn, 2017). Las cifras que se emplean para conocer los cálculos de la brecha salarial son: Salario, Cantidad bruta, Cantidades comparables y Cantidades agrupadas.

Una de las manifestaciones más visibles de la brecha salarial ocurre cuando, ante condiciones laborales similares, hombres y mujeres reciben remuneraciones distintas. Cadena (2020) explica que la discriminación salarial acontece cuando un individuo con similares características a otro, y que solo difiere en sexo, religión u otras características, recibe menor salario por razones inexplicables distintas al desempeño en su trabajo.

Históricamente, las mujeres han asumido de manera desproporcionada las responsabilidades de cuidado y trabajo no remunerado, lo que ha limitado su continuidad laboral y sus oportunidades de desarrollo profesional. Esta situación afecta su acumulación de capital humano y, en consecuencia, sus niveles de ingreso (Goldin, 2014; OCDE, 2021). A pesar de la mejoría, mayor presencia a nivel educativo y participación laboral de las mujeres, persisten diferencias significativas en los ingresos entre hombres y mujeres.

Las mujeres han experimentado transformaciones en su vida, como una disminución en el número de hijos, un aumento en el nivel educativo y una mayor integración en el mercado laboral. Sin embargo, la brecha salarial se debe en gran medida a la discriminación y a la estructura salarial vigente, donde las mujeres tienden a ocupar trabajos de menor calificación y, por ende, con salarios más bajos. A pesar de los avances en educación y participación laboral femenina, persisten diferencias significativas en los ingresos entre hombres y mujeres. Estas desigualdades no pueden explicarse únicamente por diferencias en capital humano, sino que también responden a factores estructurales y discriminación en el mercado laboral (ONU Mujeres, 2023; Blau y Kahn, 2017).

La discriminación salarial es menos pronunciada en áreas urbanas, posiblemente debido a la creciente educación de las mujeres, pero sigue siendo relevante en sectores como el manufacturero y de servicios. En contextos rurales, la brecha salarial tiende a ser más pronunciada, lo que refleja mayores limitaciones estructurales y desigualdades en el acceso a oportunidades laborales (Banco Mundial, 2020).

Desde la perspectiva de la economía feminista, se analizan los mecanismos mediante los cuales el mercado laboral reproduce desigualdades de género, evidenciando cómo factores estructurales y sociales influyen en la participación y remuneración de las mujeres (Rodríguez, 2015). Estas ideologías aspiran a construir “una sociedad más justa que reconozca y garantice la igualdad plena y efectiva de todos los seres humanos”. La economista Corina Rodríguez argumenta que dentro de este marco de pensamiento se desarrolla el concepto de economía feminista, que analiza la participación económica de las mujeres y “revela los mecanismos de discriminación dentro del mercado laboral”, ayudando a comprender cómo las estructuras y prácticas del mercado perpetúan las desigualdades de género.

Estas distintas teorías aportan perspectivas diversas sobre cómo se determinan los salarios y la existencia de una brecha salarial de género. En el marco de esta investigación, se reconoce que una parte de la brecha salarial puede explicarse por diferencias en capital humano; sin embargo, también se consideran los argumentos teóricos que destacan la existencia de retornos diferenciados y discriminación de género como factores relevantes en la determinación de los ingresos. Sin embargo, la discriminación de género sigue siendo un objetivo significativo dentro de las causas de la disparidad salarial.

La discriminación de género en el ámbito laboral constituye uno de los principales factores asociados a la persistencia de la brecha salarial. Desde esta perspectiva, las mujeres pueden enfrentar barreras para acceder a puestos de mayor jerarquía, oportunidades de desarrollo profesional y condiciones laborales equivalentes a las de los hombres, aun cuando posean

niveles similares de educación, experiencia o productividad. Estas desigualdades limitan sus oportunidades de crecimiento profesional y contribuyen a la persistencia de diferencias en las remuneraciones.

La literatura especializada sostiene que la brecha salarial de género responde a un fenómeno multidimensional en el que intervienen factores económicos, sociales, culturales e institucionales. En este sentido, la desigualdad salarial no puede atribuirse exclusivamente a las características individuales de las personas trabajadoras, sino también a las estructuras del mercado laboral y a los mecanismos de discriminación que condicionan el acceso, la permanencia y la promoción dentro del empleo (Banco Mundial, 2020; OCDE, 2021; ONU Mujeres, 2025).

En consecuencia, el análisis de la brecha salarial requiere un enfoque integral que considere tanto las diferencias en capital humano como los factores estructurales asociados a la discriminación de género, los cuales pueden influir en la forma en que el mercado remunera las características de hombres y mujeres.

1.6 Análisis histórico de los salarios y la brecha salarial de género

La evolución de los salarios y de la brecha salarial de género ha estado estrechamente vinculada a las transformaciones económicas, sociales e institucionales experimentadas por los distintos países a lo largo del tiempo. Los cambios en los modelos productivos, las políticas laborales, el acceso a la educación y la participación de las mujeres en el mercado de trabajo han influido tanto en la determinación de los salarios como en las diferencias de ingresos entre hombres y mujeres. Por ello, el análisis histórico permite comprender cómo estos factores han contribuido a configurar las desigualdades salariales observadas en la actualidad.

Los niveles salariales presentan una importante heterogeneidad derivada de diversos factores económicos y laborales. Entre los principales determinantes se encuentran el nivel educativo, la experiencia laboral, el tipo de contratación, la actividad económica desempeñada y la ubicación geográfica, elementos que influyen de manera significativa en las remuneraciones

percibidas por los trabajadores (OCDE, 2022). Asimismo, el salario puede variar dependiendo de la metodología utilizada para su estimación y de la fuente de información empleada, por lo que su análisis requiere considerar el contexto y los criterios bajo los cuales se construyen los distintos indicadores salariales.

Uno de los factores que históricamente ha tenido mayor incidencia sobre el ingreso es la educación. Desde la teoría del capital humano, una mayor escolaridad suele asociarse con mejores oportunidades laborales y mayores niveles salariales. Diversos estudios han señalado que el retorno económico de la educación tiende a incrementarse conforme aumenta el nivel académico alcanzado, lo que evidencia que la formación constituye una ventaja importante dentro del mercado laboral. Sin embargo, dichos retornos pueden variar según el contexto económico, el sector productivo y las características específicas del empleo (OCDE, 2025).

A partir de estos elementos, la literatura ha documentado que las diferencias salariales no solo responden a características productivas de los trabajadores, sino también a factores estructurales relacionados con la organización de los mercados laborales y las oportunidades de acceso al empleo. En este contexto surge el interés por estudiar la evolución histórica de los salarios y, particularmente, la persistencia de las brechas salariales entre hombres y mujeres.

El análisis de los salarios ha sido un tema central en la economía laboral, dado que refleja la interacción entre el capital y el trabajo, así como las condiciones sociales y económicas que determinan la distribución del ingreso (Smith, 1776; Marx, 1867).

Uno de los fenómenos más persistentes en los estudios económicos y sociológicos es la desigualdad salarial. En México, este tema resulta fundamental para comprender las dinámicas del mercado laboral y las condiciones que enfrentan las mujeres en distintos niveles de ingreso (INEGI, 2023).

Arceo-Gómez y Campos-Vázquez (2018) analizan la evolución de la brecha salarial de género en México entre 1990 y 2010 utilizando datos censales, encontrando cambios

importantes en la distribución de la desigualdad salarial a lo largo del tiempo. A través de la evaluación de estos datos, los autores buscan examinar la evolución que ha trascendido durante este periodo de tiempo acerca de la disparidad salarial de género.

En la literatura económica, destacan conceptos como el *sticky floor* (piso pegajoso), que describe las dificultades de las mujeres para salir de los niveles salariales más bajos, y el *glass ceiling* (techo de cristal), que hace referencia a las barreras invisibles que limitan su acceso a puestos de alta remuneración (Albrecht et al., 2003).

En el caso de México, la evidencia muestra que la brecha salarial de género ha disminuido gradualmente en los últimos años; sin embargo, esta reducción no ha sido homogénea. Arceo Y Campos (2018) señalan que la brecha ha cambiado su distribución, concentrándose en mayor medida en los niveles salariales más bajos, mientras que en los niveles más altos se ha reducido parcialmente.

Sin embargo, esta reducción no fue homogénea para todos los niveles de ingreso ni para todas las ocupaciones. De hecho, al analizar la distribución salarial en su totalidad, se encuentra que las mujeres en los niveles salariales más bajos siguen enfrentando una desigualdad fundamental, mientras que en los niveles más altos la brecha se ha reducido considerablemente.

El análisis histórico de los salarios y la brecha salarial en México revela una evolución compleja, marcada por periodos de estancamiento y recientes esfuerzos por mejorar el poder adquisitivo de los trabajadores. Durante décadas, el salario mínimo en México permaneció prácticamente congelado, con incrementos poco significativos que no alcanzaban a compensar los efectos de la inflación. Esto provocó una pérdida constante del poder de compra para millones de trabajadores, lo cual se reflejaba en niveles altos de pobreza laboral y desigualdad.

A partir de 2018, se implementaron incrementos sostenidos al salario mínimo en México, lo

que permitió una recuperación significativa del poder adquisitivo de los trabajadores. En este contexto, el salario mínimo ha experimentado aumentos importantes en términos reales, alcanzando niveles considerablemente superiores a los observados en décadas anteriores (CONASAMI, 2024).

Estos incrementos han continuado en años recientes como parte de una política sostenida de mejora salarial. Este incremento acumulado permitió que más trabajadores accedieran a la canasta básica, lo que significó un avance concreto en la reducción de la pobreza sin generar presiones inflacionarias relevantes, contrario a las advertencias de ciertos sectores empresariales.

Aunque estas medidas representan un logro importante, la brecha salarial entre hombres y mujeres continúa siendo una problemática estructural en México. A pesar de que las mujeres constituyen una parte creciente de la fuerza laboral, sus ingresos son significativamente inferiores. En México, las mujeres perciben ingresos significativamente menores que los hombres, con brechas que varían dependiendo del sector, nivel educativo y tipo de ocupación (INEGI, 2023).

Datos recientes muestran que, incluso con niveles educativos similares, las mujeres continúan percibiendo menores ingresos que los hombres (INEGI, 2023). Esta desigualdad persiste incluso en niveles altos de escolaridad, lo que indica que la brecha salarial no puede explicarse únicamente por diferencias en capital humano. Diversos estudios han documentado que, aun con características similares, las mujeres enfrentan menores retornos a la educación y a la experiencia laboral en comparación con los hombres (Blau Y Kahn, 2017). Asimismo, factores como la maternidad y las responsabilidades de cuidado han sido identificados como elementos que afectan negativamente la trayectoria laboral y los ingresos de las mujeres, ampliando la brecha salarial (Goldin, 2014).

No obstante, algunos avances se han registrado en sectores específicos. Entre 2018 y 2022, la brecha de género en el salario por hora para las mujeres asalariadas de menores ingresos

se redujo de 14.1 % a 7.5 %, lo que indica que las políticas de incremento salarial también han tenido cierto efecto positivo sobre las mujeres que se encuentran en la base de la pirámide laboral. Sin embargo, estos avances aún son limitados frente a la magnitud del problema.

Otro componente crítico en el análisis de la desigualdad salarial es el tamaño de la economía informal en México. La informalidad laboral representa un factor clave en la desigualdad salarial en México, ya que una proporción significativa de la población ocupada trabaja en condiciones sin acceso a seguridad social ni prestaciones, afectando de manera desproporcionada a las mujeres (OIT, 2018; INEGI, 2023).

A pesar de los avances observados en la recuperación del salario mínimo durante los últimos años, la desigualdad salarial continúa mostrando diferencias importantes entre grupos poblacionales. Diversos estudios han señalado que factores como la segregación ocupacional, la distribución desigual de las responsabilidades de cuidado y la elevada participación de las mujeres en empleos informales contribuyen a explicar la persistencia de las brechas salariales en México. En este sentido, la evolución reciente de los salarios evidencia que los incrementos generales en las remuneraciones no necesariamente se traducen en una reducción automática de las desigualdades de género, por lo que resulta necesario considerar elementos estructurales del mercado laboral para comprender la permanencia de estas diferencias (Banco Mundial, 2020; OCDE, 2021).

El análisis desarrollado a lo largo de este capítulo permite establecer un marco teórico integral para comprender la determinación de los salarios y la existencia de la brecha salarial de género. En primer lugar, se abordó el concepto de salario desde una perspectiva económica, identificando sus principales componentes y los factores que influyen en su determinación, tales como la oferta y demanda de trabajo, la productividad y las condiciones del mercado laboral. Este enfoque permitió entender que el salario no es únicamente el resultado de una transacción económica, sino también una expresión de las dinámicas sociales y estructurales que caracterizan al mercado laboral.

Posteriormente, desde la perspectiva microeconómica, se explica cómo el salario se configura como el precio del trabajo, determinado en función del equilibrio entre la oferta y la demanda laboral, así como por la productividad marginal del trabajador. En este contexto, se introdujo la teoría del capital humano como uno de los principales marcos explicativos de las diferencias salariales, al sostener que la educación, la experiencia y las habilidades individuales influyen directamente en la productividad y, por ende, en los niveles de ingreso (Becker, 1964; Mincer, 1974).

Sin embargo, el análisis también permite identificar las limitaciones de este enfoque, particularmente en lo que respecta a la explicación de las desigualdades salariales entre hombres y mujeres. Si bien el capital humano constituye un factor relevante, no resulta suficiente para explicar la persistencia de la brecha salarial de género, lo que evidencia la necesidad de incorporar perspectivas adicionales que consideren factores estructurales y sociales.

En este sentido, se analiza la brecha salarial desde una perspectiva de género, identificando elementos como la discriminación, la segregación ocupacional, los roles de género y la distribución desigual de las responsabilidades de cuidado como factores que influyen en la desigualdad de ingresos. Estos elementos permiten comprender que la brecha salarial no es únicamente un fenómeno económico, sino también un reflejo de estructuras sociales que condicionan la participación y remuneración de las mujeres en el mercado laboral (Blau Y Kahn, 2017; Goldin, 2014).

Asimismo, el análisis histórico permite observar la evolución de los salarios y de la brecha salarial en México, destacando tanto los avances logrados en términos de incremento salarial como las persistentes desigualdades estructurales, particularmente en contextos de informalidad laboral y segmentación del mercado de trabajo. Este recorrido evidenció que, si bien se han implementado políticas orientadas a mejorar las condiciones salariales, la desigualdad de género continúa siendo un reto significativo.

En conjunto, los elementos teóricos presentados en este capítulo permiten concluir que la diferenciación salarial responde a un entramado complejo de factores, en el que interactúan tanto características individuales —como el capital humano— como elementos estructurales —como la discriminación y la segmentación del mercado laboral—. Esta dualidad resulta fundamental para el desarrollo de la presente investigación, ya que permite establecer una base sólida para el análisis de la brecha salarial de género.

En este contexto, el siguiente capítulo se enfoca en la construcción del marco metodológico de la investigación, en el cual se describirá la fuente de información utilizada, la definición de variables y la especificación del modelo econométrico. En particular, se emplea la descomposición Oaxaca-Blinder como herramienta principal para analizar la brecha salarial, lo que permitirá distinguir entre la parte explicada por diferencias en capital humano y aquella asociada a los retornos diferenciados en el mercado laboral.

La revisión teórica permitió identificar que las diferencias salariales pueden explicarse tanto por características asociadas al capital humano como por factores relacionados con la discriminación y las condiciones estructurales del mercado laboral. Estos elementos constituyen la base conceptual sobre la cual se desarrolla la presente investigación. Una vez definidos los fundamentos teóricos, resulta necesario establecer las herramientas metodológicas y econométricas que permitirán analizar empíricamente la brecha salarial de género, aspecto que se desarrolla en el siguiente capítulo.

CAPÍTULO 2. ANÁLISIS DE REGRESIÓN CON DATOS DE CORTE TRANSVERSAL

Una vez establecidos los fundamentos teóricos de la investigación, este capítulo presenta los elementos metodológicos y econométricos que sustentan el análisis empírico. En particular, se describen los diferentes tipos de datos utilizados en econometría, los principales conceptos asociados al análisis estadístico y la regresión lineal, así como los supuestos necesarios para la estimación de modelos econométricos. Finalmente, se expone la metodología de descomposición Oaxaca-Blinder, herramienta central empleada para identificar la proporción de la brecha salarial atribuible a características observables y aquella asociada a diferencias en los retornos de dichas características.

2.1 Tipos de datos: series de tiempo, corte transversal y datos de panel

Los datos utilizados en el análisis econométrico pueden clasificarse según su estructura en series de tiempo, corte transversal y datos de panel. Esta clasificación resulta fundamental para la selección de las técnicas estadísticas y econométricas adecuadas (Wooldridge, 2010; Gujarati y Porter, 2009). En este contexto, el estudio de los datos puede apoyarse en diversos métodos y técnicas, entre las cuales destacan los enfoques univariante y multivariante, dependiendo del número de variables que se busquen analizar de manera simultánea.

El análisis multivariante permite estudiar de manera simultánea la relación entre múltiples variables, siendo una herramienta fundamental en la econometría para explicar fenómenos complejos como la determinación de los salarios (Gujarati y Porter, 2009). Existen distintas técnicas de análisis multivariante, cada una con características y usos específicos; entre las principales se encuentran la regresión múltiple, el análisis factorial, el análisis de conglomerados, el análisis de componentes principales y el análisis de correspondencia. A grandes rasgos, todas estas herramientas permiten simplificar la interpretación de datos de alta dimensionalidad, realizar predicciones y segmentar grupos dentro del conjunto principal de observaciones.

Por su parte, el análisis univariante se centra en la descripción de una sola variable, utilizando medidas de tendencia central y dispersión, así como representaciones gráficas que permiten comprender su comportamiento (Wooldridge, 2010). Entre los componentes más importantes de esta técnica se encuentran las medidas de dispersión, como la varianza, así como medidas de tendencia central, como la media, la mediana y la moda; arrojando un resumen de los datos al identificar el punto central o el valor típico de la variable. De igual manera, las representaciones gráficas contribuyen a comprender la distribución de la variable y a identificar su comportamiento general.

El análisis univariante puede aplicarse tanto a datos categóricos como numéricos. Los datos categóricos, al estar conformados por diferentes grupos o clases, suelen analizarse mediante distribuciones de frecuencia y gráficos de barras para comprender su distribución. Los datos numéricos, por su parte, pueden clasificarse en continuos y discretos: los primeros pueden tomar cualquier valor dentro de un intervalo, mientras que los segundos están conformados por valores distintos y separados. En ambos casos, la elección de las técnicas de análisis y de las visualizaciones dependerá del tipo de datos que se pretenda estudiar.

Una vez planteados estos enfoques de análisis, es importante distinguir los tipos de datos según su estructura temporal.

Las series de tiempo consisten en observaciones de una variable a lo largo del tiempo, permitiendo analizar su evolución y comportamiento dinámico (Gujarati y Porter, 2009). Los datos pueden comportarse de distintas maneras a lo largo del tiempo: pueden presentar tendencia, ciclos, movimientos aleatorios o variaciones estacionales, ya sean anuales, semestrales o de otro tipo. Las observaciones de una serie de tiempo suelen denotarse por $Y_1, Y_2, \dots, Y_t$, donde Y_t representa el valor tomado por el proceso en el instante t (Ríos, 2008).

Este tipo de datos se utiliza principalmente con fines predictivos, mediante modelos que capturan la dinámica temporal de las variables. En este tipo de análisis pueden distinguirse, de manera general, dos clases de modelos. Por un lado, se encuentran los modelos deterministas, conocidos por emplear métodos de extrapolación sencillos en los que no se hace referencia explícita a la fuente o naturaleza de la aleatoriedad subyacente en la serie. Su

relativa simplicidad suele venir acompañada de una menor precisión. Un ejemplo de ello son los modelos de promedio móvil, en los cuales el pronóstico se obtiene a partir del promedio de los “n” valores inmediatamente anteriores.

Por otro lado, están los modelos estocásticos, los cuales se basan en una descripción simplificada del proceso aleatorio subyacente en la serie. Este enfoque asume que la serie observada $Y_1, Y_2, \dots, Y_t$, se extrae de un conjunto de variables aleatorias con una distribución conjunta difícil de determinar directamente, por lo que se construyen modelos aproximados que sean funcionales para la generación de pronósticos.

Asimismo, una serie de tiempo puede descomponerse en cuatro elementos principales, que, aunque no siempre son observables de forma directa, pueden estimarse. El primero es la tendencia (T), la cual representa el comportamiento predominante de la serie y puede entenderse como el cambio en la media durante un periodo prolongado. El segundo es el ciclo (C), caracterizado por oscilaciones alrededor de la tendencia con larga duración y causas no siempre claramente identificables. El tercer componente es la estacionalidad (E), que recoge movimientos periódicos que se presentan dentro de intervalos relativamente cortos. Finalmente, el componente aleatorio (A) agrupa movimientos erráticos que no siguen un patrón específico y obedecen a causas diversas, por lo que resulta prácticamente impredecible.

Los datos de corte transversal consisten en observaciones de múltiples unidades de análisis como individuos, hogares o empresas, en un mismo periodo de tiempo, lo que permite analizar diferencias entre dichas unidades en un contexto específico (Wooldridge, 2010). No todas las observaciones corresponden necesariamente al mismo instante exacto; por ejemplo, si se realiza una encuesta de ingresos a 1,000 hogares durante el mes de marzo de 2025, aunque no todos hayan sido entrevistados el mismo día, los datos siguen considerándose de corte transversal porque reflejan la situación en un mismo periodo de tiempo (durante un mes, por ejemplo) y se asume que no hay cambios relevantes en ese intervalo.

En este tipo de datos, es común asumir que las observaciones provienen de un proceso de muestreo representativo, como ocurre en las encuestas nacionales. Este supuesto, ampliamente estudiado en los cursos introductorios de estadística, simplifica el análisis y la

estimación econométrica. Sin embargo, en la práctica existen diversas formas de muestreo y, de manera particular, en el presente trabajo se utiliza una muestra proveniente de combinaciones de muestreo, como ocurre en las encuestas nacionales.

Los datos de panel combinan la dimensión de corte transversal con la dimensión temporal, permitiendo observar a las mismas unidades a lo largo del tiempo y capturar su evolución (Wooldridge, 2010). Cuando las unidades observadas no son exactamente las mismas en todos los periodos, Gujarati y Porter (2010), así como Wooldridge (2010), se refieren a este tipo de información como datos combinados (*pooled data*), ya que las observaciones provienen de diferentes muestras de corte transversal reunidas en distintos momentos del tiempo. En aplicaciones microeconómicas, este tipo de datos suele consistir en información de individuos o empresas recopilada a través del tiempo; por ejemplo, el registro del ingreso anual de 1,000 familias durante seis años consecutivos. En ese caso, se observa cómo cambia el ingreso de las mismas unidades a través del tiempo, lo que constituye un verdadero conjunto de datos de panel. Una de las principales ventajas de utilizar datos de panel en estimaciones econométricas es que permiten aumentar la precisión de los parámetros estimados, debido al mayor número de observaciones que se obtiene al combinar la dimensión de corte transversal con la dimensión temporal. Además, una característica fundamental que distingue a los datos de panel de las simples combinaciones de cortes transversales es que, durante un intervalo de tiempo, se observa a las mismas unidades de análisis. Esto permite capturar con mayor detalle la heterogeneidad individual y la evolución de los fenómenos económicos.

2.2 Modelo de regresión clásico

El modelo de regresión lineal clásico (MRLC) constituye una de las herramientas fundamentales de la econometría, ya que permite analizar la relación entre una variable dependiente y un conjunto de variables explicativas (Gujarati y Porter, 2009; Wooldridge, 2013). Su objetivo es explicar el comportamiento de una variable dependiente en función de una o varias variables independientes. Este modelo, basado en el método de Mínimos

Cuadrados Ordinarios (MCO), permite estimar la relación entre variables y realizar inferencias estadísticas con base en ciertos supuestos teóricos.

En el análisis económico, el MRLC se utiliza ampliamente para estudiar la relación entre el salario y sus determinantes, tales como la educación, la experiencia laboral y otras características individuales. De esta manera, permite explicar diferencias en los ingresos y evaluar la importancia relativa de cada variable en su determinación.

El MRLC se representa en su forma general como:

$$Y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \cdots + \beta_k x_{ki} + u_i$$

$$\text{o en notación matricial: } Y_i = X_\beta + u$$

Donde Y es un vector columna de $n \times 1$ que representa la variable dependiente; X es la matriz de variables explicativas (de dimensiones $n \times k$), que incluye una columna de unos para el intercepto; β es un vector de parámetros desconocidos de dimensión $k \times 1$; y u es el vector de términos de error o perturbaciones aleatorias (Hayashi, 2000). El objetivo del modelo consiste en encontrar los estimadores $\hat{\beta}$ que minimicen la suma de los cuadrados de los errores, es decir, $\sum u_i^2$.

Para garantizar que los estimadores obtenidos mediante el MCO tengan propiedades deseables, el MRLC se fundamenta en una serie de supuestos considerados esenciales (Wooldridge, 2013). En primer lugar, el supuesto de linealidad en los parámetros implica que el modelo sea lineal respecto a los coeficientes β_j , aunque puede incluir transformaciones no lineales de las variables explicativas, siempre que la relación con los parámetros permanezca lineal.

Además, se asume que los datos provienen de una muestra aleatoria representativa, lo cual asegura que las observaciones sean independientes y estén idénticamente distribuidas. Este

supuesto es fundamental para la validez de los procedimientos de inferencia estadística.

Otro supuesto relevante es la ausencia de multicolinealidad perfecta entre las variables explicativas. La multicolinealidad ocurre cuando una variable independiente está altamente correlacionada con una o varias de las otras variables del modelo, lo que puede inflar las varianzas de los estimadores e impedir una interpretación confiable de los coeficientes (Gujarati y Porter, 2009). En el caso de multicolinealidad perfecta, el modelo no puede estimarse mediante MCO.

El supuesto de esperanza condicional cero del error, expresado como $E(u_i|X) = 0$, es decir, el término de error no debe estar correlacionado con el conjunto de variables explicativas. Este es uno de los supuestos más importantes, ya que garantiza que los estimadores sean insesgados (Wooldridge, 2013). Si este supuesto no se cumple, el modelo se encuentra mal especificado y los estimadores resultan sesgados e inconsistentes.

Por su parte, la homocedasticidad, otro supuesto central, implica que la varianza del término de error es constante a lo largo de todas las observaciones, es decir, $Var(u_i|X) = \sigma^2$. Si este supuesto no se cumple y existe heteroscedasticidad, los estimadores siguen siendo insesgados, pero dejan de ser eficientes, lo que afecta la validez de las pruebas de significancia.

De igual forma, el supuesto de no autocorrelación establece que los errores no deben estar correlacionados entre sí, es decir $Cov(u_i, u_j) = 0$ para $i \neq j$. Aunque este problema es más común en modelos de series de tiempo, su presencia puede afectar la eficiencia de los estimadores y la validez de las inferencias (Stock Y Watson, 2011).

Finalmente, en muchos casos se asume que los errores siguen una distribución normal, $u_i \sim N(0, \sigma^2)$. Si bien este supuesto no es estrictamente necesario para la estimación de los coeficientes mediante MCO, es fundamental para la validez de las pruebas de significancia y la construcción de intervalos de confianza, especialmente en muestras pequeñas (Greene, 2012). En muestras grandes, el teorema del límite central permite relajar esta condición, dado

que los estimadores tienden a ser aproximadamente normales.

Bajo los supuestos del MRLC, los estimadores obtenidos mediante el método de mínimos cuadrados ordinarios (MCO) presentan propiedades estadísticas deseables: son insesgados, ya que en promedio coinciden con los verdaderos parámetros poblacionales; eficientes, al tener la menor varianza entre los estimadores lineales insesgados (teorema de Gauss-Markov); y consistentes, debido a que convergen en probabilidad a los verdaderos valores de los parámetros conforme aumenta el tamaño de la muestra (Wooldridge, 2013).

El modelo de regresión también permite realizar inferencias sobre los parámetros estimados. A través de pruebas *t* se puede determinar la significancia individual de los coeficientes, mientras que, mediante la prueba *F* se puede evaluar la significancia conjunta de un grupo de variables. Asimismo, los intervalos de confianza permiten establecer rangos probables para los parámetros con cierto nivel de confianza, y el modelo puede ser utilizado para hacer predicciones sobre la variable dependiente, tanto dentro de la muestra como para observaciones nuevas fuera de la muestra, siempre y cuando se cumplan los supuestos establecidos (Gujarati y Porter, 2009).

Una característica importante del MRLC es su sensibilidad a la especificación del modelo. La omisión de variables relevantes puede generar sesgo en los estimadores, mientras que la inclusión de variables irrelevantes aumenta la varianza de estos. Por ello, la selección adecuada de variables resulta fundamental (Stock Y Watson, 2011); y suele basarse tanto en la teoría económica como en evidencia empírica previa.

Asimismo, la presencia de valores atípicos o influyentes puede afectar significativamente los resultados del modelo, razón por la cual es recomendable realizar análisis de diagnóstico como los residuos estandarizados, a través de pruebas gráficas que permitan identificar observaciones que se desvíen considerablemente del comportamiento general de los datos. Una vez que se identifican valores fuera del promedio, se debe identificar si se trata de un error en la muestra o si existió algún evento que haya hecho que cambiara abruptamente el comportamiento.

En la práctica, se utilizan diversas pruebas para verificar el cumplimiento de los supuestos del MRLC. Por ejemplo, las pruebas de Breusch-Pagan o White se utilizan para detectar heteroscedasticidad, mientras que los tests de Durbin-Watson o Breusch-Godfrey permiten evaluar la autocorrelación. En caso de detectarse violaciones a los supuestos, pueden aplicarse métodos alternativos de estimación como los Mínimos Cuadrados Generalizados (MCG) o los errores estándar robustos, con el fin de obtener inferencias válidas en presencia de heteroscedasticidad o autocorrelación (Greene, 2012).

En términos estadísticos, el MRLC también es valioso porque sirve como punto de partida para el desarrollo de modelos econométricos más complejos, como los modelos de regresión múltiple con variables ficticias, interacciones, variables instrumentales y modelos de datos de panel o series de tiempo. Su comprensión resulta, por tanto, fundamental para el análisis empírico en economía.

En el presente estudio, el modelo de regresión lineal clásico constituye la base metodológica para analizar la relación entre el salario de los individuos y sus características observables. En particular, permite evaluar el impacto de variables como la educación, la experiencia laboral y otras características sobre los ingresos.

Asimismo, este modelo sirve como fundamento para la aplicación de la descomposición Oaxaca-Blinder, la cual permite separar la brecha salarial en una parte explicada por diferencias en características observables y otra asociada a diferencias en los retornos de dichas características.

2.3 Variables binarias y categóricas

En el ámbito de la econometría, la inclusión y manejo adecuado de variables cualitativas, particularmente las variables binarias y categóricas, resultan fundamentales para capturar efectos que no pueden ser representados únicamente por variables cuantitativas. Su incorporación permite construir modelos más completos y realistas, capaces de reflejar con mayor precisión las relaciones entre variables económicas (Wooldridge, 2013; Gujarati y

Porter, 2009).

En la práctica econométrica, es bastante común trabajar con variables explicativas, e incluso dependientes, que no son numéricas continuas, sino cualitativas. En estos casos, el uso de variables binarias y categóricas permite adaptar el modelo a contextos más amplios y realistas, incorporando características individuales, sociales o institucionales que influyen en el fenómeno analizado (Wooldridge, 2013).

Las variables binarias, también conocidas como variables “dummy”, se utilizan para representar atributos cualitativos que únicamente pueden asumir dos estados posibles, como sexo (hombre o mujer), propiedad de vivienda (propietario o no), condición de empleo (empleado o desempleado), entre otros. Estas variables se codifican generalmente con los valores cero y uno, donde el valor uno indica la presencia de una característica específica, mientras que el valor cero representa su ausencia. Su inclusión en el modelo permite medir el efecto diferencial de dicha característica sobre la variable dependiente.

De acuerdo con Wooldridge (2013), el uso adecuado de variables dummy facilita una interpretación clara de los efectos marginales, al tiempo que permite incorporar información cualitativa sin alterar la linealidad del modelo en los parámetros. En este sentido, el coeficiente asociado a una variable binaria puede interpretarse como la diferencia en el valor esperado de la variable dependiente entre el grupo que presenta la característica (valor 1) y el grupo de referencia (valor 0), manteniendo constantes las demás variables del modelo.

Cuando una variable categórica tiene más de dos categorías, como el nivel educativo o región geográfica, es necesario transformarla en un conjunto de variables binarias. En este caso, se generan $k - 1$ variables dummy, donde k representa el número de categorías, con el fin de evitar la multicolinealidad perfecta, conocida como la “trampa de la variable dummy” (Gujarati y Porter, 2009). Por ejemplo, si una variable educativa incluye categorías como “preparatoria”, “universitaria” y “maestría”, se pueden crear variables dummy para dos de ellas, tomando la tercera como categoría base. Los coeficientes de estas variables dummy representan la diferencia en la variable dependiente respecto a la categoría base.

Esta técnica, conocida como codificación de referencia, es esencial para garantizar la correcta estimación del modelo. Cada coeficiente asociado a una variable dummy representa el efecto diferencial respecto a la categoría omitida, la cual actúa como punto de comparación. Es importante señalar que la elección de la categoría base no afecta la calidad del ajuste del modelo ni sus predicciones, pero sí influye en la interpretación de los coeficientes (Wooldridge, 2013).

Wooldridge (2013) enfatiza que las variables dummy no sólo son útiles en modelos de regresión con variables explicativas cualitativas, sino también en la creación de modelos con interacciones. Esto significa que se pueden evaluar efectos diferenciados entre grupos, por ejemplo, es posible evaluar si el retorno a la educación difiere entre hombres y mujeres al incluir un término de interacción entre una variable de escolaridad y una variable dummy de género. De esta manera, se pueden capturar efectos diferenciados entre grupos, enriqueciendo el análisis econométrico.

Cuando la variable dependiente también es binaria, el modelo de regresión lineal clásico presenta limitaciones importantes (Wooldridge, 2013; Greene, 2012). Específicamente, al aplicar Mínimos Cuadrados Ordinarios a una variable dependiente que sólo puede tomar los valores 0 y 1, se generan predicciones fuera del intervalo $[0,1]$, lo cual carece de interpretación probabilística. Además, la varianza del error en este caso se vuelve heteroscedástica por construcción, ya que depende de la media condicional. Para abordar este tipo de situaciones, Wooldridge (2013) propone el uso de modelos de respuesta cualitativa como el Modelo Logit y el Modelo Probit. Ambos se basan en una función de distribución acumulativa para asegurar que las probabilidades predichas se mantengan en el intervalo válido y se estiman mediante el método de Máxima Verosimilitud.

El Modelo Logit asume una distribución logística acumulativa para modelar la probabilidad de que la variable dependiente sea igual a 1, dada una combinación lineal de variables explicativas. Matemáticamente, se expresa como:

$$P(Y = 1 | X) = \frac{e^{X\beta}}{1 + e^{X\beta}}$$

Por su parte, el Modelo Probit utiliza la función de distribución acumulativa de la normal estándar para definir esa misma probabilidad:

$$P(y = 1 | X) = \Phi(X\beta)$$

En ambos casos, la interpretación directa de los coeficientes no es tan inmediata como en el modelo lineal, debido a la naturaleza no lineal de estas funciones. No obstante, es posible calcular los efectos marginales para conocer cómo una unidad de cambio en una variable explicativa afecta la probabilidad de éxito, manteniendo constantes las demás variables. Wooldridge señala que estos efectos marginales pueden evaluarse en el promedio de las variables, para obtener una medida representativa del impacto esperado.

El uso de variables cualitativas también se extiende a modelos más complejos, como los modelos de datos de panel o modelos con efectos fijos y aleatorios. En estos casos, las variables dummy permiten controlar por efectos no observables que son constantes en el tiempo para cada unidad de análisis. Así, en modelos de efectos fijos, la inclusión de variables dummy por individuo o entidad permite eliminar el sesgo de variables omitidas inobservables, siempre que estas no varíen a lo largo del tiempo. Wooldridge (2013) desarrolla ampliamente esta aplicación y destaca su utilidad en estudios longitudinales y comparativos (Wooldridge, 2013).

En aplicaciones empíricas, las variables dummy son ampliamente utilizadas para evaluar diferencias entre grupos poblacionales, analizar el impacto de políticas públicas, diferenciar el comportamiento de distintos sectores económicos, o estudiar el efecto de características demográficas como el consumo, el empleo, la inversión o el ingreso (Gujarati y Porter, 2009). Asimismo, su correcta especificación exige no sólo su inclusión en el modelo, sino también la realización de pruebas de hipótesis adecuadas para verificar si sus efectos son estadísticamente significativos. Las pruebas t para coeficientes individuales y la prueba F conjunta para grupos de dummies permiten evaluar su relevancia dentro del modelo

(Wooldridge, 2013).

El tratamiento adecuado de variables binarias y categóricas en econometría implica la adaptación e incorporación de variables cualitativas en modelos econométricos. Su correcta especificación e interpretación requiere comprensión conceptual, habilidad matemática y criterio empírico, y representa una extensión natural del modelo de regresión lineal clásico hacia un marco más general y flexible. Estos enfoques permiten analizar su impacto en las variables de interés, enriqueciendo así el análisis económico y proporcionando una comprensión más completa de los fenómenos estudiados.

En el contexto del presente estudio, el uso de variables binarias resulta fundamental para capturar diferencias entre grupos, particularmente en el análisis de la brecha salarial entre hombres y mujeres. En específico, la variable de género permite identificar si existen diferencias en los ingresos que no pueden explicarse únicamente por características observables como la educación o la experiencia laboral. De esta manera, la inclusión de variables cualitativas en el modelo econométrico contribuye a un análisis más completo y preciso de las desigualdades salariales (Blau Y Kahn, 2017).

2.4 Modelos con interacciones entre variables cualitativas

En la econometría moderna el tratamiento de variables cualitativas representa un componente central en la estimación y análisis de modelos de regresión (Wooldridge, 2013; Gujarati y Porter, 2009). El enfoque tradicional de inclusión de variables cualitativas en modelos lineales supone que los efectos de cada variable dummy son aditivos, es decir que se considera que el impacto de una característica cualitativa específica sobre la variable dependiente es constante, independientemente del valor que tomen otras variables cualitativas o cuantitativas en el modelo (Wooldridge, 2013).

No obstante, esta suposición puede resultar restrictiva en muchos contextos empíricos, especialmente cuando se sospecha que la relación entre una variable explicativa y la variable dependiente cambia según la presencia o ausencia de otra característica. En estos casos, es necesario considerar la posibilidad de interacción entre variables cualitativas, las cuales permiten modelar efectos conjuntos y capturar de manera más precisa la heterogeneidad entre

los individuos (Wooldridge, 2013; Greene, 2012).

De acuerdo con Wooldridge (2013), la inclusión de interacciones entre variables cualitativas permite que el efecto de una variable sobre la variable dependiente dependa del valor que tome otra variable. En términos prácticos, esto se implementa mediante la creación de una nueva variable que resulta del producto entre dos variables dummy. Por ejemplo, si se cuenta con una variable dummy para género (D_1 , igual a 1 si el individuo es mujer y 0 si es hombre) y otra para una característica adicional (D_2), entonces el término $D_1 \times D_2$ permitirá capturar el efecto conjunto de ambas condiciones.

Para ilustrar este concepto, puede considerarse un modelo en el que el logaritmo del salario se explica en función de la educación, el género, la pertenencia a un grupo específico y la interacción entre estas variables. El modelo puede expresarse como:

$$\ln(wage_i) = \beta_0 + \beta_1 female_i + \beta_2 nonwhite_i + \beta_3 (female_i \times nonwhite_i) + \beta_4 educ_i + u_i$$

En este caso, el coeficiente β_1 mide el efecto promedio de ser mujer sobre el log del salario para el grupo base, β_2 representa el efecto de pertenecer al segundo grupo para los hombres mientras que β_3 captura el efecto adicional asociado a la combinación de ambas características. Es decir, no se trata simplemente de sumar los efectos independientes, sino de modelar directamente el efecto conjunto. Si β_3 es estadísticamente significativo, eso implica que el efecto conjunto de las dos condiciones es distinto del efecto que se obtendría si se sumaran sus contribuciones por separado.

Desde una perspectiva de interpretación económica, este tipo de especificación resulta especialmente relevante, ya que permite identificar con mayor precisión grupos de población que podrían enfrentar barreras específicas o condiciones diferenciadas en el mercado laboral. Sin considerar interacciones, estos efectos podrían pasar desapercibidos o interpretarse de manera incompleta (Blau Y Kahn, 2017).

Una cuestión metodológica clave que resalta Wooldridge es que al incluir términos de

interacción entre variables dummy, es fundamental mantener las variables originales en el modelo. Es decir, no es suficiente incluir únicamente el producto de las variables, sino también deben incorporarse cada dummy individualmente, para garantizar una correcta interpretación de los coeficientes y evitar problemas de especificación (Wooldridge, 2013). Si se omiten las variables individuales, la interpretación del modelo se distorsiona, y los efectos no pueden separarse adecuadamente.

Además de las interacciones entre variables cualitativas, también es común considerar interacciones entre variables cualitativas y cuantitativas. Por ejemplo, podría explorarse si el efecto de la educación sobre el salario difiere entre hombres y mujeres, mediante la inclusión de un término de interacción entre la variable de escolaridad y una dummy de género. El modelo puede representarse de la siguiente manera:

$$\ln(wage_i) = \beta_0 + \beta_1 female_i + \beta_2 educ_i + \beta_3(female_i \times educ_i) + u_i$$

En este caso, el coeficiente β_3 indica en qué medida el efecto de la educación sobre el salario difiere entre hombres y mujeres, permitiendo capturar diferencias en los retornos del capital humano entre ambos grupos (Mincer, 1974; Wooldridge, 2013).

El uso de interacciones entre variables aporta mayor flexibilidad en los modelos, pero también exigen precaución en su especificación. Si se incluyen interacciones de manera excesiva, especialmente cuando el tamaño de la muestra es limitado, pueden surgir problemas de multicolinealidad o sobreajuste, lo cual reduce la precisión de las estimaciones (Gujarati y Porter, 2009). Por ello, su incorporación debe estar justificada tanto por la teoría económica como por la evidencia empírica.

Asimismo, las interacciones pueden extenderse a variables categóricas con múltiples niveles. En estos casos, se generan variables dummy para cada categoría (excepto la base) y se construyen términos de interacción con otras variables. Sin embargo, este procedimiento incrementa rápidamente el número de variables en el modelo, por lo que es necesario evaluar cuidadosamente su inclusión para evitar complejidad innecesaria (Wooldridge, 2013).

En términos generales, los modelos con interacciones entre variables cualitativas constituyen una extensión natural del modelo de regresión lineal clásico, permitiendo capturar efectos heterogéneos y representar de manera más realista las relaciones económicas. Esta herramienta resulta particularmente útil en el análisis de desigualdades, discriminación y evaluación de políticas públicas, donde distintas combinaciones de características pueden generar resultados diferenciados (Blau Y Kahn, 2017).

En el contexto del presente estudio, las interacciones adquieren especial relevancia, ya que permiten analizar si los retornos a variables como la educación o la experiencia laboral difieren entre hombres y mujeres. Este enfoque es fundamental para comprender la naturaleza de la brecha salarial, ya que no solo permite identificar diferencias en características observables, sino también en la forma en que el mercado remunera dichas características (Blau Y Kahn, 2017; Goldin, 2014).

De esta manera, el análisis de interacciones se relaciona directamente con metodologías como la descomposición Oaxaca-Blinder, en la cual las diferencias en los coeficientes estimados entre grupos reflejan precisamente diferencias en los retornos a las características individuales. Por tanto, la inclusión de interacciones y el estudio de efectos heterogéneos constituyen un paso clave para el análisis econométrico de la brecha salarial (Oaxaca, 1973; Blinder, 1973).

2.5 Bondad de ajuste

En el análisis econométrico, una vez estimado un modelo de regresión, resulta fundamental evaluar qué tan bien dicho modelo logra explicar el comportamiento de la variable dependiente. Para ello, se emplean diversas medidas de bondad de ajuste, las cuales permiten cuantificar la calidad del modelo y su capacidad explicativa (Gujarati y Porter, 2009; Wooldridge, 2013).

Una de las medidas más utilizadas es el coeficiente de determinación, denotado como R^2 , el cual indica la proporción de la variabilidad total de la variable dependiente que es explicada

por las variables independientes incluidas en el modelo. Su valor se encuentra acotado entre 0 y 1, donde valores cercanos a 1 indican un mayor poder explicativo del modelo, mientras que valores cercanos a 0 sugieren una capacidad limitada para explicar la variación observada en la variable dependiente (Gujarati y Porter, 2009).

El coeficiente de determinación se define como:

$$R^2 = 1 - \frac{\sum u_i^2}{\sum (Y_i - \bar{Y})^2}$$

donde $\sum u_i^2$ representa la suma de los cuadrados de los residuos y $\sum (Y_i - \bar{Y})^2$ corresponde a la suma total de cuadrados. Esta medida permite evaluar qué tan bien se ajusta el modelo a los datos observados.

Sin embargo, el R^2 presenta una limitación importante, ya que su valor tiende a incrementarse al agregar nuevas variables explicativas, incluso si estas no resultan relevantes para explicar la variable dependiente. Como consecuencia, puede generar una percepción engañosa de mejora en el modelo (Wooldridge, 2013).

Para corregir esta limitación, se utiliza el coeficiente de determinación ajustado, denotado como $R_{ajustado}^2$, el cual penaliza la inclusión de variables innecesarias. Este indicador toma en cuenta el número de variables explicativas y el tamaño de la muestra, proporcionando una medida más adecuada para comparar modelos con distinto número de regresores (Gujarati y Porter, 2009).

El $R_{ajustado}^2$ se define como:

$$R_{ajustado}^2 = 1 - \left(\frac{n-1}{n-k-1} \right) (1 - R^2)$$

donde n es el número de observaciones y k el número de variables explicativas del modelo. Además del R^2 , existen otros criterios que permiten evaluar la calidad de un modelo econométrico, especialmente cuando se comparan distintas especificaciones. Entre estos se encuentran los criterios de información, como el Criterio de Información de Akaike (AIC, por sus siglas en inglés) y el Criterio de Información Bayesiano (BIC, por sus siglas en

inglés). Ambos criterios incorporan una penalización por la complejidad del modelo, favoreciendo aquellos que logran un buen ajuste con un menor número de parámetros.

El AIC se define como:

$$AIC = -2\ln(L) + 2k$$

mientras que el BIC se expresa como:

$$BIC = -2\ln(L) + k\ln(n)$$

donde L representa la función de verosimilitud del modelo, k el número de parámetros estimados y n el tamaño de la muestra. En ambos casos, valores más bajos indican un mejor equilibrio entre ajuste y parsimonia (Greene, 2012).

Otra herramienta utilizada en la evaluación de modelos es la prueba de Kolmogorov-Smirnov (K-S), la cual permite comparar la distribución empírica de los datos con una distribución teórica específica. Esta prueba resulta útil para evaluar la normalidad de los residuos, lo cual es relevante en el contexto de la inferencia estadística (Greene, 2012).

Asimismo, es posible analizar la bondad de ajuste mediante el estudio de los residuos del modelo, los cuales representan la diferencia entre los valores observados y los valores estimados. El análisis gráfico de los residuos permite detectar patrones que podrían indicar problemas de especificación, como la presencia de heterocedasticidad, autocorrelación o no linealidad (Wooldridge, 2013).

En este sentido, un buen modelo econométrico no solo debe presentar un alto poder explicativo, sino también cumplir con los supuestos del modelo de regresión lineal clásico. Por ello, la evaluación de la bondad de ajuste debe complementarse con pruebas de diagnóstico que permitan validar la consistencia de los resultados obtenidos.

En el contexto del presente estudio, estas medidas permiten evaluar la capacidad del modelo para explicar las variaciones en el salario de los individuos, así como comparar distintas especificaciones del modelo econométrico utilizadas para analizar la brecha salarial entre hombres y mujeres. De esta manera, la bondad de ajuste no solo constituye un criterio

estadístico, sino también una herramienta fundamental para seleccionar el modelo más adecuado para el análisis empírico (Stock Y Watson, 2011).

2.6 Pruebas de diagnóstico

Las pruebas de diagnóstico en regresión con datos de corte transversal constituyen herramientas estadísticas utilizadas para analizar la validez de los supuestos del modelo de regresión. El objetivo principal de estas pruebas es verificar si se satisfacen las condiciones necesarias para que los resultados estimados sean confiables, consistentes y estadísticamente interpretables (Wooldridge, 2013; Stock Y Watson, 2011). En un análisis de regresión con datos de corte transversal, donde la información es recolectada en un único punto en el tiempo o en un periodo corto, estas pruebas adquieren una importancia fundamental, ya que el análisis no involucra la dinámica temporal propia de los modelos de series de tiempo, sino la comparación entre distintas unidades de observación.

Las pruebas de diagnóstico permiten examinar si el modelo econométrico cumple con los supuestos fundamentales del modelo de regresión lineal clásico. Si estos supuestos no se satisfacen, las estimaciones obtenidas mediante Mínimos Cuadrados Ordinarios pueden perder propiedades deseables, lo que afecta tanto la precisión de los coeficientes como la validez de las pruebas de hipótesis e intervalos de confianza (Gujarati y Porter, 2009).

El primer supuesto hace referencia a la linealidad de la función, es decir, a que la relación entre la variable dependiente Y y las variables independientes $X_1, X_2, \dots, X_k$ debe expresarse de forma lineal en los parámetros. Matemáticamente, este supuesto se representa como:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k + \varepsilon$$

Si este supuesto no se cumple, los coeficientes estimados pueden resultar sesgados o el modelo puede no representar adecuadamente la relación económica entre las variables. Aunque en algunos casos es posible incorporar transformaciones no lineales de las variables, como cuadrados, logaritmos o interacciones, la linealidad en los parámetros debe mantenerse para que el modelo siga siendo estimable bajo el enfoque clásico (Wooldridge, 2013).

Otro aspecto esencial dentro de las pruebas de diagnóstico es la correcta especificación de las variables incluidas en el modelo. La omisión de variables relevantes o la inclusión de variables irrelevantes puede generar errores de especificación, afectando la interpretación de los resultados. En este sentido, las pruebas de diagnóstico no solo sirven para verificar propiedades estadísticas del error, sino también para evaluar si el modelo se encuentra bien planteado en términos de la relación entre las variables explicativas y la variable dependiente (Stock Y Watson, 2011).

Uno de los problemas más importantes en este tipo de análisis es la heterocedasticidad, la cual se presenta cuando la varianza del término de error no permanece constante entre las observaciones. Este fenómeno es particularmente frecuente en datos de corte transversal, ya que individuos, hogares o empresas con diferentes características pueden exhibir distintos niveles de variabilidad en sus errores. En presencia de heterocedasticidad, los estimadores de Mínimos Cuadrados Ordinarios continúan siendo insesgados, pero dejan de ser eficientes, lo que afecta la precisión de las estimaciones y puede invalidar las pruebas de significancia (Gujarati y Porter, 2009; Wooldridge, 2013).

Para detectar la heterocedasticidad, suelen emplearse pruebas como la de Breusch-Pagan y la de White. Estas pruebas permiten evaluar si la varianza de los errores depende de las variables explicativas del modelo. Cuando se detecta este problema, es posible recurrir a errores estándar robustos u otras correcciones que permitan mantener la validez de la inferencia estadística (Greene, 2012).

Otro aspecto relevante en las pruebas de diagnóstico es la normalidad de los errores. Aunque este supuesto no es estrictamente necesario para estimar los coeficientes del modelo mediante Mínimos Cuadrados Ordinarios, sí resulta importante para la realización de pruebas de hipótesis y la construcción de intervalos de confianza, especialmente cuando el tamaño de la muestra no es suficientemente grande. Si los residuos se apartan considerablemente de una distribución normal, las inferencias basadas en pruebas t y F pueden perder precisión (Wooldridge, 2013).

Para evaluar la normalidad de los residuos, es posible utilizar pruebas estadísticas como la

de Kolmogorov-Smirnov o la de Shapiro-Wilk, así como herramientas gráficas como histogramas y gráficos de probabilidad normal. Estas técnicas permiten analizar si la distribución de los residuos se aproxima razonablemente a una distribución normal, lo cual es deseable para fortalecer la confiabilidad del modelo (Greene, 2012).

Asimismo, el análisis de los residuos constituye una parte central del diagnóstico econométrico. Los residuos representan la diferencia entre los valores observados y los valores estimados por el modelo, y su comportamiento puede revelar diversos problemas de especificación. Por ejemplo, patrones sistemáticos en los residuos pueden indicar relaciones no lineales omitidas, variables relevantes no incluidas o la presencia de valores atípicos que influyen de manera importante en los resultados (Gujarati y Porter, 2009).

Otro elemento que puede evaluarse dentro del diagnóstico del modelo es la multicolinealidad, que ocurre cuando dos o más variables explicativas están altamente correlacionadas entre sí. Aunque este problema no sesga los coeficientes estimados, sí puede inflar sus errores estándar y dificultar la interpretación individual de los efectos de cada variable. En la práctica, una herramienta ampliamente utilizada para detectar este problema es el Factor de Inflación de la Varianza (VIF) (Wooldridge, 2013).

De igual manera, pueden aplicarse pruebas orientadas a verificar la correcta especificación funcional del modelo, como la prueba Ramsey RESET, la cual permite detectar si existen variables omitidas, formas funcionales inadecuadas o errores en la construcción general del modelo (Stock Y Watson, 2011). Este tipo de pruebas resulta útil porque permite identificar si la estructura planteada por el investigador representa de forma adecuada el fenómeno económico analizado.

En este sentido, las pruebas de diagnóstico no solo permiten verificar el cumplimiento de los supuestos teóricos del modelo de regresión, sino que también ayudan a fortalecer su especificación y su consistencia empírica. Cuando se detectan problemas, el investigador puede recurrir a distintas soluciones, como la transformación de variables, la reespecificación del modelo, el uso de errores estándar robustos o la exclusión de observaciones atípicas justificadas.

En el contexto del presente estudio, se pone especial énfasis en la evaluación de la heterocedasticidad y la normalidad de los residuos, debido a su importancia para la validez de los resultados obtenidos en los modelos estimados. Estas pruebas resultan particularmente relevantes porque permiten asegurar que las estimaciones econométricas, así como la posterior aplicación del método Oaxaca-Blinder, se realicen bajo condiciones que favorezcan inferencias confiables.

De esta manera, la aplicación de pruebas de diagnóstico constituye una etapa esencial del análisis econométrico, ya que permite validar los resultados obtenidos, identificar posibles deficiencias en la especificación del modelo y fortalecer la interpretación de la brecha salarial (Wooldridge, 2013), entre hombres y mujeres dentro del sector analizado.

Como complemento a las pruebas de diagnóstico descritas anteriormente, la Figura 2 presenta un ejemplo de implementación en el software R para la estimación del modelo de regresión y la aplicación de pruebas de heterocedasticidad, multicolinealidad, normalidad de los residuos y especificación del modelo. Este tipo de herramientas permite llevar a la práctica los conceptos teóricos abordados y facilita la validación empírica de los supuestos del modelo econométrico.

Figura 2. Ejemplo de implementación en R

```
# Ajustar el modelo de regresión
modelo <- lm(y ~ x1 + x2, data = datos)
# Prueba de heterocedasticidad (Breusch-Pagan)
library(lmtest)
bptest(modelo)
# Comprobar multicolinealidad (VIF)
library(car)
vif(modelo)
# Prueba de normalidad de los residuos (Shapiro-wilk)
shapiro.test(residuals(modelo))
# Prueba de especificación del modelo (Ramsey RESET)
resettest(modelo)
```

Fuente: Elaboración propia.

2.7 Técnica econométrica Oaxaca-Blinder

El método Oaxaca-Blinder constituye una de las técnicas econométricas más utilizadas en el análisis de la brecha salarial, particularmente en estudios que buscan explicar las diferencias

de ingreso entre distintos grupos de la población. Este enfoque fue desarrollado por Ronald Oaxaca (1973), quien propuso una metodología para descomponer las diferencias salariales entre hombres y mujeres en componentes explicables y no explicables. Posteriormente, Alan Blinder (1973) amplió esta técnica mediante su aplicación empírica a datos reales, consolidando así la metodología conocida como descomposición Oaxaca-Blinder.

Esta técnica tiene como objetivo principal descomponer una variable (en este caso, el salario) en dos grupos, permitiendo identificar los factores que explican las diferencias observadas entre ellos. En su planteamiento original, la descomposición Oaxaca-Blinder establece que la diferencia salarial entre dos grupos se debe a dos componentes fundamentales: por un lado, las diferencias en capital humano, consideradas explicables, que se derivan de variables observables como la escolaridad o la experiencia laboral; y por otro, un componente no explicado, que suele asociarse con discriminación u otros factores no observables. Esta distinción resulta fundamental, ya que permite no solo identificar la existencia de una brecha salarial, sino también comprender sus posibles causas y cuantificar el peso relativo de cada uno de sus componentes (Oaxaca, 1973; Blinder, 1973).

En este sentido, el método Oaxaca-Blinder es ampliamente utilizado en investigaciones económicas y sociales, especialmente en el ámbito laboral, para analizar la brecha salarial entre géneros, así como entre otros grupos como trabajadores nativos e inmigrantes. Su principal aporte radica en que permite no solo identificar la existencia de una diferencia salarial, sino también descomponerla en factores observables, como educación, experiencia laboral, edad y sector laboral; y factores no observables, como discriminación, habilidades innatas o preferencias individuales; proporcionando así una estimación cuantitativa de la contribución de cada uno de estos elementos a la brecha salarial (Blau Y Kahn, 2017).

Desde el punto de vista teórico, esta metodología se sustenta en la teoría del capital humano, desarrollada por Jacob Mincer (1974), quien propone un modelo en el que el salario se explica en función de la escolaridad y la experiencia laboral. De acuerdo con este enfoque, la acumulación de capital humano influye directamente en el nivel de ingresos de los

individuos, siendo los coeficientes de la ecuación una medida del impacto de estas variables sobre el salario (Oaxaca, 1973).

La ecuación minceriana se representa de la siguiente manera:

$$\ln(y) = \beta_0 + \beta_1 S + \beta_2 X + \beta_3 X^2 + \varepsilon$$

Donde:

$\ln(y)$ es el ingreso,

S es la variable que mide el nivel de escolaridad,

X corresponde a la experiencia laboral,

X^2 es el término cuadrático de la experiencia que permite capturar rendimientos decrecientes,

β_0 es la constante del modelo,

β_1 y β_2 son los coeficientes asociados a las variables explicativas, y

ε es el término de error o perturbación.

A partir de esta especificación, es posible estimar ecuaciones salariales diferenciadas para cada grupo de análisis. En el caso de la brecha salarial por género, estima ecuaciones con salarios de hombres (h) y salario de mujeres (m) por lo tanto, estas ecuaciones se expresan de la siguiente manera:

$$Y_h = X'_h \beta_h + U_h$$

$$Y_m = X'_m \beta_m + U_m$$

donde Y_h representa el ingreso de los hombres y Y_m el ingreso de las mujeres, X'_h y X'_m corresponden a los vectores de variables independientes para cada grupo; y U_h y U_m representan los términos de perturbación asociados a cada ecuación.

Con base en estas estimaciones, la metodología Oaxaca-Blinder permite descomponer la diferencia en los salarios promedio entre hombres y mujeres mediante la siguiente expresión:

$$\bar{Y}_h - \bar{Y}_m = (\bar{X}_h - \bar{X}_m) \hat{\beta}_h + (\hat{\beta}_h - \hat{\beta}_m) \bar{X}_m$$

Donde:

$\bar{Y}_h - \bar{Y}_m$ es la diferencia en los salarios promedio de hombres y mujeres

$(\bar{X}'_h - X'_m)\hat{\beta}_h$ representa las diferencias en dotación de capital humano, es decir, las variables observables que distinguen a hombres y mujeres.

$(\hat{\beta}_h - \hat{\beta}_m)X'_m$ corresponde al componente no explicado, la distinción entre coeficientes de hombres y mujeres, el cual se interpreta como la discriminación existente o como el efecto de factores no observables (Blinder, 1973; Blau Y Kahn, 2017).

De manera complementaria, el método también puede representarse mediante una medida relativa de discriminación salarial, que compara los salarios observados con aquellos que se obtendrían en ausencia de discriminación (Oaxaca, 1973):

$$D = \frac{\frac{w_m}{w_f} - \left(\frac{w_m}{w_f}\right)^0}{\left(\frac{w_m}{w_f}\right)^0}$$

Donde:

$\frac{w_m}{w_f}$ simboliza la relación entre el salario observado para hombres y para mujeres (Wage male y Wage female).

Mientras que $\left(\frac{w_m}{w_f}\right)^0$ representa la relación salarial, en un escenario hipotético sin discriminación. Por lo que es útil para evaluar el grado de desigualdad salarial atribuible a factores distintos al capital humano.

No obstante, esta metodología también presenta algunas limitaciones que es importante considerar en la interpretación de los resultados (Jann, 2008). En primer lugar, la descomposición depende de la forma en que se especifica el modelo, por lo que la selección de variables y la forma funcional pueden influir directamente en los resultados obtenidos. Asimismo, el método requiere contar con información suficientemente detallada y consistente, ya que cualquier variable relevante que no sea incluida en el modelo se traslada

al componente no explicado.

En este sentido, el componente no explicado debe interpretarse con cautela, ya que, si bien suele asociarse con la existencia de discriminación, en realidad puede estar capturando también otros factores no observados, como diferencias en habilidades, trayectorias laborales o condiciones específicas del empleo que no se encuentran disponibles en la base de datos. Finalmente, si bien la literatura ha propuesto diversas extensiones del modelo (como su aplicación en regresiones cuantílicas o en modelos no lineales), en este trabajo se adopta la versión clásica de Oaxaca-Blinder, al considerarse adecuada para los objetivos del estudio y para el tipo de información disponible.

En resumen, el método Oaxaca-Blinder constituye una herramienta econométrica sólida para el análisis de la brecha salarial, ya que permite no solo identificar la existencia de diferencias en los ingresos entre grupos, sino también descomponer dichas diferencias en función de sus principales determinantes. A partir de esta metodología, es posible distinguir entre el papel de las características observables (como la escolaridad o la experiencia laboral) y la influencia de factores no observados en la generación de desigualdades salariales.

Con base en los elementos teóricos y metodológicos desarrollados a lo largo de este capítulo, en el siguiente apartado se procede a la aplicación empírica del modelo, utilizando información de la Encuesta Nacional de Ocupación y Empleo (ENOE), con el propósito de analizar la brecha salarial entre hombres y mujeres dentro de Banca y Seguros. De esta manera, se busca trasladar el marco conceptual presentado hacia un análisis cuantitativo que permita evaluar, en términos concretos, la magnitud y las fuentes de dicha desigualdad (Oaxaca, 1973; Blinder, 1973).

Los conceptos y herramientas metodológicas desarrollados en este capítulo proporcionan el marco analítico necesario para la estimación de los modelos econométricos utilizados en la investigación. Una vez definida la metodología, resulta indispensable presentar la fuente de información, el proceso de construcción de la base de datos y las características generales de la población analizada, aspectos que se abordan en el siguiente capítulo.

CAPÍTULO 3. CONSTRUCCIÓN E INTERPRETACIÓN DE LA BASE DE DATOS PARA MÉXICO Y PARA EL CASO DE ESTUDIO

En este capítulo se presenta la fuente de información utilizada en la investigación, así como el proceso de depuración y construcción de la base de datos analítica. A partir de la Encuesta Nacional de Ocupación y Empleo (ENOE), correspondiente al primer trimestre de 2025, se realiza una caracterización descriptiva de la población estudiada mediante el análisis de variables sociodemográficas, laborales y salariales. Este análisis permite identificar patrones generales en la distribución de los ingresos y establecer un primer acercamiento a las diferencias salariales observadas entre hombres y mujeres.

3.1 México: análisis histórico de los salarios

El análisis histórico de los salarios en México expone una evolución ardua hacia la mejora del bienestar económico de la población trabajadora, influida por los diversos factores económicos, sociales y políticos en México, que se han presentado a lo largo del tiempo (CONASAMI, 2024; INEGI, 2023). En este capítulo se revisan en términos generales los aspectos económicos y sociales, como las condiciones laborales de épocas diferentes de México, para comprender mejor el fenómeno.

En la primera mitad del siglo XX, y producto de la revolución mexicana, inicia la implementación de salarios mínimos, gracias a las conquistas de los trabajadores, incorporándolo en la constitución nacional, con la finalidad de evitar que los salarios mínimos cayeran por debajo de ese nivel inferior establecido, necesario para satisfacer sus necesidades y las de sus familias (CPEUM, Artículo 123). En el Porfiriato en el periodo de 1877-1911 existió un crecimiento económico importante. La economía durante el Porfiriato presentó un importante crecimiento económico acompañado de elevados niveles de desigualdad social.

En 1917 se promulgó la Constitución Política de los Estados Unidos Mexicanos, considerada una legislación de vanguardia a nivel mundial. En el contexto de la Revolución Mexicana,

cuando la estructura económica y social del país se encontraba dominada por un sistema capitalista y persistían las desigualdades heredadas del Porfiriato, uno de sus principales objetivos fue establecer un nuevo marco jurídico que garantizara derechos sociales y evitara el retorno a las condiciones prevalecientes durante ese periodo. Con esto se le garantizaba que México pudiera entrar en una etapa de industrialización, en la que, con mayores salarios, la fuerza de trabajo pudiera ser más productiva. Actualmente, se establece en la Carta Magna que el salario mínimo “deberá ser suficiente para satisfacer las necesidades normales de un jefe de familia en el orden material, social y cultural, y para proveer la educación obligatoria de los hijos” (CPEUM, Artículo 123, apartado VI)

La OIT (Organización Internacional del Trabajo) nos dice que, el salario mínimo debe ser “la suma mínima que deberá pagarse al trabajador por el trabajo o servicios prestados, dentro de un lapso determinado, bajo cualquier forma que sea calculado, por hora o por rendimiento, que no puede ser disminuida ni por acuerdo individual, ni colectivo, que está garantizada por ley y puede fijarse para cubrir las necesidades mínimas del trabajador y de su familia, teniendo en consideración las condiciones económicas y sociales de los países.” (OIT, 1992). Es por ello por lo que, se pone en marcha en diversos países durante la primera mitad del siglo XX y surgen legislaciones. En México, la legislación en la materia tuvo un objetivo social para reducir la pobreza y la desigualdad mediante aumentos significativos al salario mínimo (Marinakís, 2006). En México, como en algunos países de América Latina, la forma de asegurar esos salarios es mediante la CONASAMI (Comisión Nacional de Salarios Mínimos), dependiente del Poder Ejecutivo.

En el tiempo de la industrialización, el salario en México tuvo puntos de quiebre y de recuperación. Mientras que durante la etapa de posguerra y los primeros años del proceso de industrialización el salario mínimo experimentó periodos de pérdida de poder adquisitivo, a partir de la década de 1940 comenzó una etapa de crecimiento económico sostenido conocida como el Desarrollo Estabilizador o Milagro Mexicano (1940-1970). Este modelo económico se caracterizó por la promoción de la industrialización por sustitución de importaciones, la estabilidad de precios, el crecimiento sostenido del producto interno bruto y una mayor intervención del Estado en la economía. En este contexto, los salarios reales mostraron una

tendencia favorable y el poder adquisitivo de los trabajadores registró una recuperación gradual. Posteriormente, a partir de 1951, el salario mínimo inició un ascenso continuo que se prolongó hasta 1976, año en el que alcanzó uno de sus niveles históricos más altos en términos reales (Moreno-Brid y Ros, 2009). Posteriormente, las políticas económicas implementadas durante los gobiernos de Luis Echeverría (1970- 1976) y José López Portillo (1976-1982), se inicia una disminución que finaliza en la crisis de los 80, con la principal caída del salario mínimo real de todo el siglo XX.

A principios de la década de los ochentas, en Mexico y parte de las naciones latinoamericanas ponen en marcha el modelo conservador neoliberal, en un entorno mundial con debilitamiento del bloque socialista y surgimiento de un contexto internacional caracterizado por cambios en los modelos económicos y una reducción en la intervención del Estado en la economía, teniendo en cuenta esto, en México comienza una época de crisis, ya que familias adineradas se beneficiaron, mientras que una gran parte de la población cayó en condiciones de mayor empobrecimiento. Durante esta década, fue el escenario del mayor desplome en términos reales del salario. En la crisis de los 80s, el salario mínimo real, tuvo una caída aproximadamente del 65 % en su poder adquisitivo (CONASAMI, 2024), considerando como punto de partida 1976.

Posteriormente, inicia una crisis conocida como “error de diciembre”, del año 1994, frase acuñada por el expresidente mexicano Carlos Salinas de Gortari, la cual inició por la falta de reservas internacionales generando la devaluación del peso mexicano y un aumento significativo en la inflación. Trayendo como consecuencia que el salario mínimo real haya sufrido la caída adicional del 10 % en un solo año, con ello incrementando la pérdida de poder adquisitivo que ya se venía arrastrando desde la década anterior. La crisis económica de 1994 que vivió México fue iniciada en los primeros días de la presidencia de Ernesto Zedillo (Banco de México, 1995).

A partir de la crisis de la deuda de 1982, México inició un proceso de transformación económica caracterizado por la adopción de políticas orientadas a la estabilidad macroeconómica, la apertura comercial y la reducción de la intervención del Estado en la economía. En este contexto, el salario mínimo dejó de concebirse principalmente como un

instrumento para mejorar el bienestar de los trabajadores y pasó a utilizarse como un mecanismo de control inflacionario. Como consecuencia, el crecimiento de los salarios quedó subordinado a los objetivos de estabilidad económica, lo que provocó un deterioro sostenido de su poder adquisitivo durante las décadas siguientes (Esquivel, 2015; CONASAMI, 2024).

Durante el sexenio de Miguel de la Madrid (1982-1988) se implementaron programas de ajuste estructural como respuesta a la crisis económica y al elevado endeudamiento externo. Las políticas de austeridad y control de precios priorizaron la estabilización de la economía; sin embargo, ello implicó una importante pérdida del poder adquisitivo de los salarios. Posteriormente, durante la administración de Carlos Salinas de Gortari (1988-1994) se promovió la apertura comercial, la privatización de empresas públicas y la firma del Tratado de Libre Comercio de América del Norte (TLCAN). Aunque estas medidas buscaban impulsar el crecimiento económico y mejorar la competitividad del país, la política salarial continuó orientada a contener la inflación, por lo que los incrementos al salario mínimo permanecieron limitados y no lograron recuperar el poder de compra perdido.

La crisis financiera de 1994 marcó el inicio del sexenio de Ernesto Zedillo (1994-2000), provocando un fuerte incremento de la inflación y una nueva reducción del salario real. A pesar de la posterior recuperación macroeconómica, los aumentos salariales continuaron siendo insuficientes para compensar el deterioro acumulado. De acuerdo con la CONASAMI (2024), durante este periodo el salario mínimo registró un crecimiento nominal de 148 %, mientras que la inflación acumulada alcanzó aproximadamente 225 %, lo que evidencia la pérdida sostenida del poder adquisitivo.

Durante las administraciones de Vicente Fox (2000-2006), Felipe Calderón (2006-2012) y Enrique Peña Nieto (2012-2018) se mantuvo una política económica enfocada principalmente en preservar la estabilidad de precios y el equilibrio macroeconómico. Aunque existieron expectativas de mejorar las condiciones laborales y elevar los ingresos de la población, los incrementos al salario mínimo continuaron siendo moderados y, en diversos periodos, apenas compensaron la inflación. Como resultado, el salario mínimo permaneció durante varios años por debajo de los niveles necesarios para satisfacer las necesidades

básicas de los trabajadores, contribuyendo al incremento de la desigualdad y al crecimiento del empleo informal (Esquivel, 2015; CONASAMI, 2024).

Entre 1995 y 2018, el salario mínimo experimentó el periodo más prolongado de deterioro de su poder adquisitivo. Diversos estudios estiman que perdió entre 70 % y 75 % de su capacidad de compra debido a que fue utilizado como un instrumento de política antiinflacionaria, limitando de manera sistemática sus incrementos reales (CONASAMI, 2024; Esquivel, 2015). Esta estrategia permitió mantener la estabilidad de precios, pero redujo significativamente la capacidad del salario mínimo para cumplir con su función como mecanismo de protección del ingreso de los trabajadores.

Un cambio importante ocurrió en 2019, con el inicio de la administración del presidente Andrés Manuel López Obrador, cuando se implementó la Nueva Política de Salarios Mínimos, cuyo principal objetivo fue recuperar el poder adquisitivo perdido durante las décadas anteriores. Esta política impulsó incrementos reales al salario mínimo y estableció la Zona Libre de la Frontera Norte, donde se fijó un salario mínimo superior al del resto del país para fortalecer la competitividad económica de esa región. A diferencia de las políticas aplicadas en décadas previas, la recuperación salarial se desarrolló sin generar presiones inflacionarias significativas, de acuerdo con los resultados reportados por la CONASAMI (2024).

Esta estrategia ha tenido continuidad durante la administración de la presidenta Claudia Sheinbaum Pardo, quien ha planteado como uno de sus compromisos incrementar progresivamente el salario mínimo hasta alcanzar un nivel equivalente a 2.5 veces la línea de pobreza por ingresos urbanos al concluir el sexenio. Para ello, se contempla mantener incrementos reales al salario mínimo con el propósito de fortalecer el ingreso de los trabajadores, reducir la pobreza laboral y consolidar el proceso de recuperación del poder adquisitivo iniciado en 2019 (CONASAMI, 2025).

En conjunto, la evolución histórica del salario en México refleja el tránsito de un largo periodo caracterizado por la contención salarial como instrumento de estabilidad macroeconómica hacia una etapa orientada a la recuperación del poder adquisitivo de los

trabajadores. Si bien los incrementos recientes han contribuido a mejorar las condiciones de ingreso y a reducir parcialmente la pobreza laboral, persisten desafíos importantes relacionados con la informalidad, la productividad y la necesidad de mantener un equilibrio entre el crecimiento de los salarios y la estabilidad de precios (Banco de México, 2024).

3.2 Aportación de las actividades económicas de Banca y Seguros

Uno de los ámbitos económicos que aportan de manera significativa en México en cuanto al crecimiento económico se encuentra las actividades bancarias y de seguros, los cuales proporcionan servicios financieros orientados a la intermediación financiera y la protección frente a riesgos económicos. Hay una vinculación estrecha entre ambas ramas de actividad en cuanto a su aportación a la economía, a la estabilidad financiera y a la protección del patrimonio tanto de personas como de empresas. Y esto a su vez promueve la productividad, el bienestar general y una mayor competitividad en México (Banco de México, 2024; CNBV, 2024).

Además de su relevancia económica, la industria bancaria y aseguradora se caracteriza por demandar personal altamente calificado, con niveles educativos superiores al promedio nacional y una amplia participación de profesionistas especializados. Estas características lo convierten en un espacio de interés para el estudio de las diferencias salariales entre hombres y mujeres, ya que permite analizar si las brechas de ingreso persisten incluso en contextos donde el capital humano es relativamente homogéneo.

La actividad bancaria en México como ya se mencionó es muy importante en el desarrollo del país, porque facilita a las y los mexicanos el acceso al financiamiento, promover la inclusión financiera como también contribuir al crecimiento sostenible de México; la cual está evolucionando y en la actualidad se ha elegido a la inteligencia artificial y el análisis de datos para mejorar la eficiencia de sus servicios (Deloitte, 2024).

La banca en México no solo tiene como función administrar el ahorro y canalizar créditos de forma eficiente, sino que otras de sus funciones es garantizar un sistema de pago sólido con esto contribuyendo a la estabilidad macroeconómica de México. En el año 2023, la actividad

bancaria en México registró resultados históricos, con un crecimiento del 15 % en utilidades, alcanzando los 273 mil millones de pesos y un ROE (Return on Equity) del 18.5 %, superior al promedio de largo plazo, según datos de Deloitte. Por lo tanto, la Cámara de Diputados dio a conocer que la inversión extranjera directa en servicios financieros teniendo lugar como la segunda más relevante del país, con 5,492.9 millones de dólares, de los cuales el 79.4 % se dirigió a la banca múltiple. Dando a conocer que las cifras muestran la confianza de inversionistas y clientes en la solidez del sector financiero mexicano (Deloitte, 2024; Cámara de Diputados, 2024).

Para el año 2024, se registró que el 47 % del crédito bancario se destinó a empresas, facilitando el crecimiento de industrias clave como también, el fortalecimiento empresarial (CNBV, 2024). Además, el tener acceso a servicios bancarios es fundamental para la inclusión financiera como también el bienestar de los mexicanos que hacen uso del sistema bancario. El *Panorama Anual de Inclusión Financiera 2024* dice que en México existían al cierre del año 2023 un total de 16,576 cuentas en instituciones financieras por cada 10,000 adultos, relevando un aumento en servicio y productos que ayuda a más personas gestionar su economía, ahorrar y planificar su futuro (CNBV, 2024). Sin embargo, en la actualidad México enfrenta desafíos en términos de inclusión financiera, ya que presenta altos niveles de rentabilidad bancaria, pero bajos niveles de acceso al financiamiento respecto a su PIB.

Otro lado fundamental de la banca comercial es contar con un sistema financiero sólido y estable ya que los precios de los activos financieros reflejen sus condiciones de oferta y demanda, contribuyen a una asignación eficiente de los recursos financieros entre ahorradores, deudores e inversionistas, con ello logrando que los agentes económicos continúen. Por lo tanto, permite una transformación de rendimientos, es decir que las instituciones financieras puedan transferir fondos de agentes con recursos disponibles y que realizan depósitos a corto plazo a agentes con necesidades financieras de largo plazo. Para que el sistema financiero funcione de manera correcta, el Banco de México juega un papel importante en promover la estabilidad financiera como también prevenir alguna crisis económica. En este sentido, el artículo 2 de la Ley del Banco de México establece que son también finalidades del Banco promover el sano desarrollo del sistema financiero y propiciar

el buen funcionamiento de los sistemas de pagos.

La Institución Aseguradora puede definirse como “manifestación técnica y organizada de las iniciativas socio económicas de compensación de riesgos” (Guardiola, 1990). Inicia a raíz de la existencia del riesgo y sus consecuencias negativas el cual viene acompañado de arreglar los daños económicos producidos por el evento negativo.

De acuerdo con Skipper (1997), el seguro contribuye al crecimiento económico al reducir la incertidumbre asociada a la actividad productiva, permitiendo que los individuos y las empresas asuman proyectos de inversión que, de otra manera, resultarían demasiado riesgosos. El seguro no es simplemente una característica del crecimiento económico; constituye una necesidad para la gran mayoría de las economías actuales. Lamentablemente, los vínculos precisos entre los seguros y el desarrollo económico han sido poco estudiados. Como consecuencia, la importancia de las aseguradoras para la economía suele ser subestimada por los responsables de diseñar políticas públicas.

Skipper (1997) sostiene que el seguro contribuye al desempeño económico de al menos siete formas distintas:

- a) En primer lugar, promueve la estabilidad financiera y reduce la incertidumbre. Esto permite que las empresas operen con menor volatilidad y menor riesgo de quiebra, proporcionando mayor estabilidad financiera y social a la economía. Al mismo tiempo, reduce la ansiedad tanto de las personas como de los empresarios frente a eventos adversos.
- b) Como segundo aporte, los seguros privados pueden complementar o sustituir ciertos programas de seguridad social gubernamentales. Esto reduce la presión sobre los recursos públicos y puede contribuir a una asignación más eficiente de los recursos de la sociedad.
- c) En tercer lugar, los seguros facilitan el comercio y la actividad empresarial. Las economías modernas se sustentan en la especialización productiva, la cual requiere un sistema financiero flexible y eficiente. La existencia de una amplia variedad de productos de seguros, así como la innovación constante en servicios y precios,

favorece las actividades comerciales e impulsa el desarrollo económico. En ausencia de mecanismos adecuados de aseguramiento, el comercio y la inversión pueden verse limitados por los riesgos asociados a la actividad económica.

- d) Como cuarta consideración, dicha actividad contribuye a la movilización del ahorro nacional. Los países con mayores tasas de ahorro tienden a registrar mayores tasas de crecimiento económico. Las aseguradoras, especialmente las compañías de seguros de vida, actúan como intermediarios financieros que canalizan los recursos de los hogares hacia inversiones productivas de largo plazo. A diferencia de los bancos comerciales, que suelen concentrarse en depósitos y créditos de corto plazo, las aseguradoras cuentan con obligaciones de largo plazo y flujos de recursos más estables, lo que las convierte en importantes proveedoras de financiamiento para gobiernos y empresas.
- e) Un quinto beneficio es que las aseguradoras permiten una gestión más eficiente de los riesgos mediante tres mecanismos fundamentales: la valoración del riesgo, la transformación del riesgo y la agrupación o diversificación del riesgo. A través de sus actividades de suscripción y de inversión, las aseguradoras generan información valiosa sobre los niveles de riesgo de individuos y empresas, contribuyendo a una toma de decisiones más eficiente por parte de inversionistas, acreedores y otros agentes económicos.

Adicionalmente, el seguro permite transferir riesgos asociados a pérdidas patrimoniales, responsabilidades legales, interrupciones de ingresos y otras contingencias. Las aseguradoras también facilitan la transformación de los perfiles de riesgo y liquidez de los ahorros de las personas y las empresas, adaptándolos a sus necesidades particulares.

- f) En sexto lugar, las aseguradoras y reaseguradoras tienen incentivos económicos para promover la prevención y reducción de pérdidas. Por ello, suelen ofrecer asesoría especializada en administración de riesgos, seguridad, prevención de accidentes y mitigación de daños. Estas acciones reducen las pérdidas económicas tanto para las empresas como para la sociedad en general.
- g) Finalmente, como séptimo aporte es que la actividad económica que desempeñan las

aseguradoras contribuye a una asignación más eficiente del capital dentro de la economía. Las aseguradoras recopilan grandes volúmenes de información para evaluar empresas, proyectos y administradores antes de otorgar cobertura o realizar inversiones. Estas decisiones funcionan como señales para el mercado sobre la viabilidad y calidad de determinadas organizaciones o proyectos. De esta manera, el seguro favorece que los recursos financieros se asignen hacia actividades más productivas y sostenibles, mejorando la eficiencia económica agregada.

En México, los seguros desempeñan un papel clave en el ámbito económico, ya que promueve la actividad económica al ofrecer mecanismos de protección financiera para personas, patrimonios y empresas. Además, contribuye a la estabilidad económica y brinda tranquilidad a los asegurados mediante coberturas relacionadas con vida, accidentes, enfermedades y retiro, entre otras. Asimismo, permite que las empresas sean más resilientes frente a diversos riesgos, incluyendo fenómenos naturales, interrupciones operativas o incumplimientos de pago, como ocurre con los seguros de crédito que protegen las cuentas por cobrar, uno de los activos estratégicos de las organizaciones.

De igual manera, las aseguradoras facilitan la gestión de riesgos financieros y de mercado asociados a fluctuaciones cambiarias, variaciones en los precios de materias primas, movimientos en las tasas de interés y otros factores económicos que pueden afectar la estabilidad y rentabilidad de las empresas. A través de distintos mecanismos de cobertura y administración de riesgos, contribuye a reducir la incertidumbre y fortalecer la capacidad de planeación de los agentes económicos.

“La actividad realizada por las aseguradoras es una de las más dinámicas y estratégicas para la economía. Nos permiten enfrentar eventos que, en gran medida, son impredecibles y cuyos efectos permean sobre los hogares, las empresas y los gobiernos” (Ramírez de la O, 2024).

Una de las aportaciones en las cuales contribuyen los seguros en México es en la producción nacional, en los últimos años existe un crecimiento en las aseguradoras, sin embargo, es importante aumentar su penetración en el mercado; es decir, ampliar su contribución a la economía nacional. El ámbito asegurador representa aproximadamente el 3 % del PIB de

México, una cifra que, aunque inferior al promedio de la OCDE (Organización para la Cooperación y Desarrollo Económicos) (8.9 %), muestra un crecimiento constante y refleja su importancia en la economía nacional (CNSF, 2024; OCDE, 2023)

En México la actividad aseguradora tiene un impacto favorable en el empleo. Ya que cuenta con amplia gama de servicios y productos ofrecidos por las aseguradoras, por lo tanto, buscan profesionales en áreas como actuaría, análisis de riesgos, ventas y atención al cliente. Estas oportunidades laborales contribuyen al crecimiento económico y al bienestar general de los trabajadores mexicanos (AMIS, 2024). De igual manera, el seguro fomenta otro tipo de sectores económicos, ya que las empresas aseguradas tienen una mayor capacidad para asumir riesgos y llevar a cabo actividades comerciales. Al contar con protección para sus bienes, las empresas pueden concentrarse en su crecimiento y expansión, lo que impulsa el crecimiento económico y fomenta la competitividad en el mercado.

Desde la perspectiva laboral, la Banca y Seguros presenta características particularmente relevantes para el análisis de la brecha salarial de género. Si bien se trata de un sector formal, con altos niveles de profesionalización y una participación importante de mujeres en diversas áreas administrativas, comerciales y técnicas, diversos estudios han señalado que la presencia femenina disminuye conforme aumenta el nivel jerárquico de los puestos. Esta situación puede generar diferencias en las trayectorias laborales, oportunidades de ascenso y niveles de remuneración, aun cuando hombres y mujeres cuenten con niveles similares de educación y experiencia.

La actividad económica Banca y Seguros se distingue por contar con una fuerza laboral altamente especializada y con elevados niveles de escolaridad. Profesiones como actuaría, economía, finanzas, administración, contabilidad, ingeniería financiera y análisis de datos tienen una participación importante dentro de las instituciones financieras y aseguradoras. Debido a ello, las remuneraciones suelen ubicarse por encima del promedio observado en otros sectores económicos, convirtiéndolo en un entorno particularmente adecuado para analizar la relación entre capital humano y niveles salariales.

Asimismo, se trata de un ámbito laboral caracterizado por estructuras organizacionales

jerarquizadas, donde existen trayectorias profesionales claramente definidas y posibilidades de crecimiento hacia puestos de supervisión, gerencia y dirección. En este contexto, el análisis de la brecha salarial resulta especialmente relevante, ya que permite identificar si hombres y mujeres acceden en igualdad de condiciones a las oportunidades de promoción y desarrollo profesional dentro de las organizaciones.

Diversos organismos internacionales han señalado que, aun cuando la participación femenina en el sector financiero ha aumentado durante las últimas décadas, las mujeres continúan enfrentando desafíos para alcanzar los niveles más altos de responsabilidad y liderazgo. Esta situación puede traducirse en diferencias en las trayectorias laborales y en los niveles de remuneración percibidos, incluso cuando existen características similares de educación, experiencia y capacitación profesional.

Adicionalmente, el sector financiero ha experimentado una creciente incorporación de mujeres durante las últimas décadas, particularmente en áreas administrativas, comerciales y de atención especializada. Sin embargo, distintos estudios han señalado que esta mayor participación no siempre se refleja en una representación equivalente dentro de los puestos directivos o en los niveles salariales más altos. Por ello, este ámbito laboral constituye un espacio de interés para analizar si las diferencias observadas en los ingresos responden únicamente a factores asociados al capital humano o si persisten elementos estructurales relacionados con el género.

En consecuencia, la Banca y Seguros representa un caso de estudio particularmente interesante para analizar la brecha salarial de género. Al tratarse de un entorno formal, con personal altamente calificado y condiciones laborales relativamente homogéneas, permite evaluar si las diferencias salariales observadas pueden explicarse únicamente por factores asociados al capital humano o si persisten elementos estructurales que generan retornos diferenciados entre hombres y mujeres.

Finalmente, la actividad económica bancaria y de seguros forma parte esencial del sistema financiero en México y tiene aportaciones significativas que fortalecen la economía de

México, fomentan la estabilidad financiera y promueven el bienestar social. A pesar de que operan con objetivos diferentes al final coinciden en varios aspectos fundamentales. Tanto la banca como los seguros contribuyen de manera directa al crecimiento económico de México.

La banca se dirige al ahorro hacia la inversión mediante el otorgamiento de créditos, mientras que, las aseguradoras fomentan el ahorro de largo plazo y respalda tanto a personas como a las empresas ante situaciones de riesgo. La complementariedad de ambas actividades económicas permite fortalecer la actividad económica, fomentar la estabilidad financiera y contribuir al desarrollo económico del México.

El análisis histórico de los salarios en México proporciona el contexto necesario para comprender la evolución de las remuneraciones y las políticas salariales a nivel nacional. No obstante, dichas transformaciones no se manifiestan de la misma forma en todos los sectores económicos, pues cada uno presenta dinámicas particulares relacionadas con su estructura ocupacional, productividad y composición de la fuerza laboral. En este sentido, la Banca y los Seguros constituyen un caso de estudio especialmente relevante debido a su importancia estratégica dentro del sistema financiero mexicano, su alto grado de formalización y la elevada participación de personal calificado. Estas características permiten analizar si las diferencias salariales entre hombres y mujeres persisten incluso en un entorno donde los niveles de educación y experiencia son relativamente similares. A partir de ello, en el siguiente apartado se presenta la fuente de información utilizada para el desarrollo del análisis empírico.

3.3 Fuente de información: Encuesta Nacional de Ocupación y Empleo

Desde su puesta en marcha en enero de 2005, la ENOE marcó un hito en la generación de estadísticas laborales en México. Sustituyendo los modelos previos como la ENE (Encuesta Nacional de Empleo) y la ENEU (Encuesta Nacional de Empleo Urbano), la ENOE introdujo mejoras conceptuales y fortalezas metodológicas para ofrecer datos más completos, precisos y comparables (INEGI, 2005).

La trayectoria de la ENOE se enmarca en la evolución de las encuestas laborales en México: se remonta a la ENH (Encuesta Nacional de Hogares, 1972), pasa por la ECMO (Encuesta

Continua de Mano de Obra, 1977), la ECSO (Encuesta Continua sobre Ocupación, 1979), hasta llegar a la ENEU en 1983. Esta sucesión permitió ampliar gradualmente la cobertura geográfica (inicialmente tres áreas metropolitanas, después 12, y más tarde casi todo el país), hasta garantizar representatividad por entidad federativa desde el año 2000 (INEGI, 2014). El diseño metodológico de la ENOE se fundamenta en un marco muestral derivado del Censo de Población y Vivienda 2000, actualizado periódicamente para reflejar la incorporación de viviendas nuevas o la remoción de las demolidas. La unidad de selección es la vivienda, la de observación es el hogar y sus integrantes, y el cuestionario incluye módulos sociodemográficos y de ocupación y empleo (INEGI, 2023).

Este cuestionario se distingue por contar con una estructura básica y otra ampliada, aplicándose la versión expandida durante el primer trimestre de cada año para obtener mayor detalle, y la básica el resto del año. Incluye temas como ingresos, horas trabajadas, prestaciones, búsqueda de empleo y condiciones críticas de ocupación, así como codificación de ocupaciones mediante el Sistema Nacional de Clasificación de Ocupaciones (SINCO, 2019; INEGI, 2023).

La metodología de recolección combinaba originalmente entrevistas presenciales y telefónicas durante la pandemia: se aplicó la ETOE (Encuesta Telefónica de Ocupación y Empleo) entre abril y junio de 2020, luego se inició la versión mixta “ENOE Nueva Edición (ENOEN)” hasta julio de 2022. Desde entonces, retornó a entrevistas totalmente presenciales, asegurando la comparabilidad de los indicadores (INEGI, s.f.).

La calidad de la información generada por la ENOE está respaldada por mecanismos de supervisión y control que incluyen entrevistas, procesos de precrítica sistemática y supervisión en campo mediante dispositivos móviles, lo que garantiza la consistencia y trazabilidad de los datos (INEGI, 2022). Los resultados de la encuesta se publican de manera mensual a nivel nacional y trimestral para las entidades federativas y 39 ciudades del país, proporcionando información tanto en tasas como en niveles absolutos de la Población Económicamente Activa (PEA), empleo, desocupación, subocupación, condiciones críticas de ocupación e informalidad laboral. Asimismo, el INEGI pone a disposición del público

microdatos, tabulados, manuales de usuario y herramientas interactivas para la consulta y análisis de la información (INEGI, 2023; INEGI, s.f.). Entre los principales indicadores generados por la encuesta se encuentran la Población Económicamente Activa, las tasas de participación económica, ocupación, desocupación y subocupación, así como las condiciones críticas de ocupación, que permiten evaluar de manera integral la situación y dinámica del mercado laboral mexicano.

En marzo de 2020 se inició la serie mensual de indicadores con cobertura histórica desde 2005, permitiendo análisis profundos y seguimiento a fenómenos como la crisis económica de 2008–2009 o la pandemia (INEGI, 2020). La tasa de participación mensual ha oscilado en torno al 60 %, mientras que las tasas de desocupación, subocupación e informalidad reflejan la relación con ciclos económicos; durante la crisis de 2008–2009, la tasa de desocupación abierta pasó de 3.5 % a más del 6 % en solo un año (INEGI, 2020; Banco de México, 2020).

La informalidad ha permanecido alta (por encima del 50 %) mientras que la subocupación afecta a entre el 5 % y 7 % de la población, reflejando vulnerabilidad laboral que no suele ser visible en indicadores macroeconómicos (INEGI, 2023).

La ENOE tiene relevancia crítica por varias razones: la cobertura territorial y plurinacional, porque incluye zonas urbanas y rurales, y no excluye sectores vulnerables, lo que la hace representativa para todo el país, sin sesgos geográficos (INEGI, 2014), la frecuencia mensual de la ENOE permite responder con agilidad ante eventos económicos como crisis, reformas o emergencias, lo cual es invaluable para la urgencia de las políticas públicas (INEGI, s.f.), Asimismo, la disponibilidad de indicadores detallados sobre informalidad, subocupación y condiciones críticas de ocupación facilita el diseño de políticas públicas específicas para mejorar condiciones de trabajo, ingresos y formalización laboral, la longitud histórica y comparabilidad de la ENOE permite reconstruir series y realizar análisis de largo plazo y evaluaciones históricas (INEGI, 2023). Y también su transparencia y acceso, ya que contiene microdatos, manuales y herramientas interactivas de fácil acceso, la ENOE pone la

información en manos de académicos, analistas, gobiernos y público general (INEGI, 2023).

La ENOE es crucial para entender los impactos de las crisis, como la mencionada anteriormente (2008–2009), cuando la desocupación abierta creció de 3.55 % a 6.12 %, destacando cómo el empleo rural informal absorbió parte del impacto. Esta dinámica no se detectaría con indicadores macroeconómicos generales sin el detalle que la ENOE provee (INEGI, 2020).

La Encuesta Nacional de Ocupación y Empleo (ENOE), constituye una de las fuentes más completas para analizar las condiciones laborales y sociodemográficas de la población mexicana. Su diseño se basa en un sistema modular de microdatos que permite vincular información de distintos niveles como vivienda, hogar y persona con el propósito de ofrecer una caracterización detallada del mercado laboral y del contexto en el que se desarrollan las actividades económicas del país. Esta organización facilita la construcción de indicadores, el análisis comparativo entre entidades federativas y la evaluación de tendencias laborales en distintos periodos (INEGI, 2023).

La primera base que integra la ENOE corresponde al módulo de Vivienda. Este componente contiene la información estructural y contextual de la unidad habitacional seleccionada dentro del diseño muestral. Incluye variables relacionadas con la ubicación geográfica, el estrato socioeconómico, el acceso a servicios, el tipo de construcción y el número de hogares presentes en la vivienda. Estas variables permiten comprender las condiciones materiales en las que habitan los hogares entrevistados y constituyen la base para identificar y enlazar los registros subsecuentes. La mayoría de los datos incluidos en este módulo corresponden a códigos numéricos y clasificaciones categóricas que describen las características físicas del inmueble y su entorno (INEGI, 2023).

Posteriormente se encuentra la base de Hogar, que describe la estructura interna y la composición de cada hogar que reside en la vivienda seleccionada. Este módulo incorpora información sobre el resultado de la entrevista, la fecha y modalidad de levantamiento, el tipo de informante y el número de integrantes del hogar. También recoge datos sobre la dinámica interna del hogar y eventos relacionados con el operativo de campo. Las variables

en esta base son principalmente categóricas y numéricas, y permiten identificar el universo de personas que serán analizadas en los módulos posteriores. Además, proporciona la conexión entre las características del hogar y las condiciones laborales o educativas de cada individuo (INEGI, 2023).

El módulo Sociodemográfico constituye uno de los componentes centrales de la ENOE, ya que integra la información básica de todas las personas residentes del hogar. Incluye variables como sexo, edad, parentesco con el jefe del hogar, nivel educativo, asistencia escolar, estado conyugal y condición de residencia. Un elemento fundamental de esta base es el factor de expansión asignado a cada individuo, el cual permite obtener estimaciones representativas de la población total. Las variables incluidas en este módulo combinan codificaciones categóricas, valores numéricos y clasificaciones estandarizadas que permiten caracterizar la estructura demográfica y educativa de la población entrevistada (INEGI, 2023).

La información sobre el mercado laboral se encuentra distribuida en dos módulos principales: el Cuestionario de Ocupación y Empleo Parte I (COE1) y Parte II (COE2). La primera parte del cuestionario recopila la información esencial para determinar la condición de actividad económica de la población de 12 o 15 años y más, dependiendo del año y versión del instrumento. Este módulo identifica si una persona se encuentra ocupada, desocupada o no económicamente activa; además, detalla características como la ocupación desempeñada, la rama de actividad, la posición en la ocupación, las horas trabajadas, la afiliación a seguridad social y los ingresos laborales obtenidos. Las variables incluidas en este módulo abarcan desde valores numéricos hasta categorías codificadas e información monetaria, lo que posibilita analizar tanto la calidad del empleo como las condiciones de inserción laboral de la población (INEGI, 2023).

La segunda parte del cuestionario, COE2, complementa la información laboral al profundizar en aspectos relacionados con las condiciones del empleo principal y, en su caso, del empleo secundario. Este módulo documenta datos sobre horas adicionales trabajadas, tipo de contrato, prestaciones laborales, características de la unidad económica, ingresos adicionales y condiciones que permiten evaluar la estabilidad, formalidad y diversificación laboral de los individuos. Las variables incluidas son tanto categóricas como numéricas, y amplían el panorama sobre la estructura del empleo en México y sobre las condiciones específicas bajo

las cuales las personas desempeñan sus actividades económicas (INEGI, 2023).

Aunque cada módulo cumple una función distinta, todos se encuentran articulados mediante identificadores comunes, como entidad federativa, municipio, número de vivienda, número de hogar y número de renglón. Esta estructura relacional permite vincular la información para realizar análisis integrales que conecten las dimensiones sociodemográficas, habitacionales y laborales. Asimismo, la ENOE utiliza clasificaciones estandarizadas para ocupaciones (SINCO) y actividades económicas (SCIAN), así como codificaciones uniformes para niveles educativos, lo que asegura la comparabilidad de los datos a lo largo del tiempo y entre distintas unidades de análisis (INEGI, 2023).

En conjunto, las bases de microdatos de la ENOE conforman un sistema robusto y versátil para el estudio del mercado laboral en México. Su estructura modular, la amplitud de variables que integra y la posibilidad de vincular distintos niveles de información permiten analizar desde características básicas de la población hasta condiciones laborales específicas como ingresos, jornada, estabilidad y formalidad. Por estas razones, la ENOE constituye una fuente indispensable para investigaciones económicas, sociales y actuariales enfocadas en comprender la dinámica del empleo y los factores que determinan las desigualdades laborales en el país.

Para el desarrollo de esta investigación se hace uso de la base Sociodemográfica (SDEM) de la ENOE, debido a que es el módulo que concentra la información esencial necesaria para caracterizar a la población y vincularla con sus condiciones laborales. Esta base contiene datos fundamentales como sexo, edad, nivel educativo, estado conyugal, características del hogar y condición de residencia, además del factor de expansión que permite obtener estimaciones representativas de la población. La amplitud y estructura de estas variables hacen posible identificar de manera precisa a los individuos que forman parte del mercado laboral y diferenciar grupos relevantes para el análisis.

La base Sociodemográfica constituye una fuente idónea para abordar el estudio de las brechas salariales en México, ya que ofrece una estructura rica, flexible y articulada que permite

realizar comparaciones entre entidades federativas y analizar de manera integral la relación entre las características individuales y las variaciones territoriales en las condiciones laborales y en los ingresos, proporcionando así un fundamento sólido para la investigación empírica. El uso de esta base resulta especialmente adecuado para investigaciones sobre la brecha salarial de género, ya que permite incorporar simultáneamente tanto variables demográficas como laborales, lo que facilita un análisis más completo de los factores asociados a las diferencias en ingresos entre hombres y mujeres. La calidad, consistencia y nivel de detalle de la información contenida en la ENOE garantizan que las estimaciones derivadas del análisis sean comparables, robustas y representativas a nivel nacional y estatal (INEGI, 2023).

3.4 Interpretación estadística y gráfica de variables de la Encuesta Nacional de Ocupación y Empleo, a nivel nacional

Para el desarrollo de la presente investigación se emplea la información correspondiente al primer trimestre de 2025 de la Encuesta Nacional de Ocupación y Empleo (ENOE), elaborada por el Instituto Nacional de Estadística y Geografía (INEGI). La selección de este periodo responde a que, durante el primer trimestre de cada año, el INEGI aplica el cuestionario ampliado, el cual incorpora información sociodemográfica y laboral con un mayor nivel de detalle respecto al cuestionario básico utilizado en los demás trimestres. Esta versión permite disponer de un conjunto más amplio de variables relacionadas con la educación, las características del empleo, las condiciones laborales y los ingresos, elementos indispensables para la construcción de las variables empleadas en el presente estudio y para la aplicación de la metodología Oaxaca-Blinder.

Dado que la ENOE es una encuesta probabilística por muestreo, la base de datos incorpora factores de expansión que permiten extrapolar la información observada en la muestra hacia el total de la población. En particular, la encuesta proporciona factores de expansión mensuales y trimestrales. El factor de expansión trimestral permite obtener estimaciones representativas para el conjunto del trimestre, mientras que el factor de expansión mensual posibilita realizar estimaciones correspondientes a cada uno de los meses que lo integran.

Estos factores corrigen las diferentes probabilidades de selección de las unidades muestrales y garantizan que las estimaciones reflejen adecuadamente la estructura de la población mexicana (INEGI, 2023; Lohr, 2010).

En la presente investigación, los factores de expansión se emplean en la elaboración de las estadísticas descriptivas con el propósito de obtener resultados representativos de la población ocupada en México. No obstante, para la estimación de los modelos econométricos y la aplicación de la descomposición Oaxaca-Blinder, el análisis se realiza sobre la muestra depurada de observaciones, ya que el interés principal consiste en estudiar las relaciones entre las variables que conforman el modelo dentro de la población objeto de estudio.

La información sociodemográfica proporcionada por la ENOE es utilizada con el propósito de analizar el comportamiento salarial en México. La muestra original incluye 423,302 observaciones, pero al revisar la consistencia y calidad de los datos fue necesario aplicar un proceso de filtración con el fin de garantizar que el análisis se realizara únicamente sobre personas con información suficiente y pertinente para el estudio.

Durante este proceso se identificó que una parte de los registros aparecía clasificada en la categoría de “ocupación no especificada”, la cual agrupa principalmente a menores de 0 a 12 años, así como a individuos de otras edades cuya información laboral no está declarada o que no proporcionaron datos sobre su condición de actividad, incluyendo su salario. Estos casos presentan información incompleta sobre variables fundamentales como salario, escolaridad, experiencia o jornada laboral.

Debido a las características de la información disponible en la ENOE, algunas de las variables utilizadas en el estudio requirieron ser construidas como aproximaciones (proxies). En particular, la encuesta no proporciona directamente una variable numérica continua que represente los años efectivos de escolaridad de los individuos, sino únicamente categorías asociadas al nivel educativo alcanzado. Por ello, fue necesario realizar una equivalencia en años de estudio para construir una variable de escolaridad potencial utilizada dentro del análisis econométrico.

De manera similar, la ENOE tampoco incorpora una variable que mida directamente la

trayectoria laboral acumulada o los años efectivos de experiencia laboral de los trabajadores. En consecuencia, la experiencia laboral fue aproximada mediante una medida de experiencia laboral potencial, calculada a partir de la expresión: (*edad – años de escolaridad – 6*). Esta aproximación busca representar el tiempo que un individuo podría haber permanecido dentro del mercado laboral después de concluir su formación académica y constituye uno de los principales componentes de los modelos de capital humano utilizados para explicar las diferencias salariales (Mincer, 1974; Becker, 1964).

Debido a ello, y considerando que el objetivo del presente estudio es analizar la brecha salarial entre trabajadores de 23 a 27 años, remunerados y pertenecientes a Banca y Seguros, estos registros fueron excluidos de la base analítica.

A diferencia de otros estudios que eliminan a los trabajadores no remunerados, en este caso se decidió conservar a los individuos clasificados como ocupados, aun cuando reciban o no remuneración monetaria, siempre y cuando su condición laboral estuviera claramente identificada y la información sobre sus características fuera completa y coherente. Esta decisión se debe a que dichos registros no presentaban inconsistencias en su clasificación de ocupación y aportan información válida sobre el comportamiento de la población trabajadora.

Después de retirar exclusivamente los casos pertenecientes a la categoría de “ocupación no especificada”, la muestra final quedó conformada por 190,503 observaciones válidas, correspondientes a personas en edad laboral y con información suficiente para estimar las variables relevantes del estudio. Esta depuración permite asegurar que el análisis de salarios y la posterior aplicación del método Oaxaca-Blinder se realicen sobre una base sólida, evitando sesgos derivados de registros incompletos o de individuos que no forman parte de la fuerza laboral.

Una vez definida la muestra final, resultó necesario analizar la composición de la población analizada con el fin de comprender la estructura general de los datos y establecer un primer acercamiento a las diferencias salariales entre hombres y mujeres. En este sentido, la Tabla 2 presenta la distribución por sexo de las personas incluidas en la muestra depurada a nivel nacional. Esta caracterización inicial permite observar la proporción de hombres y mujeres

dentro de la fuerza laboral considerada y constituye el punto de partida para el análisis descriptivo de las principales variables sociodemográficas y salariales que se desarrolla en las siguientes secciones.

Tabla 2. Distribución por sexo de la muestra depurada (ENOE 2025-I, nivel nacional)

SEXO	OBSERVACIONES	PORCENTAJE
Mujeres	78,977	41.46%
Hombres	111,526	58.54%
Total	190,503	100.00%

Fuente. Elaboración propia con base en datos de la ENOE 2025-I.

La Tabla 2 presenta la distribución por sexo de las personas incluidas en la muestra depurada a nivel nacional. Esta caracterización inicial permite identificar la composición básica de la población analizada y constituye el punto de partida para el examen descriptivo de las variables sociodemográficas y laborales. La muestra presenta una mayor participación de hombres (59 %) en comparación con las mujeres (41 %), lo cual es consistente con la estructura general del mercado laboral mexicano reportada por la ENOE.

Si bien esta distribución permite dimensionar la presencia relativa de cada grupo dentro de la población ocupada, por sí sola no ofrece información sobre sus condiciones laborales ni sobre el comportamiento de sus ingresos. Por ello, resulta necesario examinar de manera más detallada las características generales de la muestra, lo cual se desarrolla en la siguiente tabla mediante el análisis de estadísticos descriptivos agregados.

Con el fin de describir de manera integral las características sociodemográficas y laborales de la muestra, la Tabla 3. presenta los estadísticos descriptivos básicos de las principales variables utilizadas en el análisis. Esta tabla resume la distribución general de la edad, escolaridad, experiencia laboral, salario mensual y horas trabajadas, lo que permite obtener una primera aproximación al comportamiento de la población ocupada a nivel nacional.

Tabla 3. Estadísticos descriptivos generales de la muestra depurada (ENOE 2025-I, nivel nacional)

VARIABLE	MEDIA	MEDIANA	DESVIACIÓN ESTÁNDAR	MÍNIMO	MÁXIMO	OBSERVACIONES
Edad (años)	40	39	14	12	98	190,503
Escolaridad	11	12	4	0	21	185,933
Experiencia laboral en años (proxy)	22	21	15	0	86	190,503
Salario mensual (MXN)	11,608	8,364	6,784	0	50,184	156,996
Horas trabajadas por semana	39	44	16	0	60	188,308

Fuente. Elaboración propia con base en datos de la ENOE 2025-I. La variable de años de escolaridad fue construida mediante una equivalencia en años a partir del nivel educativo reportado en la encuesta. A partir de esta variable, la experiencia laboral potencial fue estimada como: $(edad - años\ de\ escolaridad - 6)$. Los valores no especificados fueron excluidos de los cálculos.

Los estadísticos descriptivos presentados en la Tabla 3 permite caracterizar de manera general a la población ocupada analizada a nivel nacional. La edad promedio de la muestra es de 40 años, con una mediana de 39, lo que indica una concentración en población adulta en edad productiva. La escolaridad promedio es de 11 años, mientras que la mediana alcanza 12 años, lo cual sugiere que una proporción relevante de la población ha completado al menos el nivel medio superior. Asimismo, la experiencia laboral potencial presenta una media de 29 años, reflejando trayectorias laborales extensas y heterogéneas.

En cuanto al salario mensual, se observa una media de \$11,608 pesos y una mediana de \$8,364 pesos, lo que confirma la presencia de una distribución asimétrica hacia la derecha, característica común en los mercados laborales donde una fracción reducida de trabajadores concentra los ingresos más altos. Esta dispersión también se evidencia en la diferencia entre el mínimo y el máximo salarial.

No obstante, estos resultados corresponden a la población ocupada en su conjunto y no permiten identificar si el comportamiento salarial es homogéneo entre hombres y mujeres.

Dado que el objetivo central del estudio es analizar la brecha salarial de género, resulta necesario desagregar los indicadores de ingreso por sexo, lo cual se presenta en la siguiente tabla.

Tabla 4. Salario mensual por sexo (ENOE 2025-I, nivel nacional)

SEXO	MEDIA	MEDIANA	DESVIACIÓN ESTÁNDAR	MÍNIMO	MÁXIMO	OBSERVACIONES
Mujeres	10,619	8,364	5,893	0	50,184	66,006
Hombres	12,325	12,546	7,279	0	50,184	90,990
Brecha salarial	13.84%					

Fuente. Elaboración propia con base en datos de la ENOE 2025-I. La brecha salarial se calculó como la diferencia relativa del salario promedio femenino respecto al masculino. $((Media_{hombres} - Media_{mujeres} / Media_{hombres}) \times 100.)$

La Tabla 4 presenta los estadísticos descriptivos del salario mensual desagregados por sexo a nivel nacional. Se observa que el ingreso promedio de los hombres asciende a \$12,325 pesos, mientras que el de las mujeres es de \$10,619 pesos. Esta diferencia representa una brecha salarial descriptiva de 13.84 %, lo que indica que, en promedio, las mujeres perciben un ingreso menor al de los hombres en el mercado laboral mexicano.

La mediana salarial confirma esta disparidad. Mientras que la mediana de los hombres alcanza \$12,546 pesos, en el caso de las mujeres se ubica en \$8,364 pesos. La diferencia entre ambas medidas sugiere que la brecha no se limita únicamente a valores extremos, sino que se mantiene también en el centro de la distribución salarial.

Asimismo, la desviación estándar es mayor en los hombres (\$7,279) que en las mujeres (\$5,893), lo que indica una mayor dispersión de ingresos dentro del grupo masculino. Esto podría estar asociado a una mayor presencia de hombres en los niveles salariales más altos o en ocupaciones con mayor variabilidad de ingresos.

En conjunto, estos resultados evidencian la existencia de diferencias salariales entre hombres

y mujeres a nivel nacional. No obstante, esta comparación es puramente descriptiva y no considera aún posibles diferencias en variables como educación, experiencia laboral o características del empleo, por lo que resulta necesario profundizar en el análisis de estos factores, lo cual se aborda en las siguientes tablas.

Si bien la Tabla 4 evidencia la existencia de una brecha salarial descriptiva entre hombres y mujeres, esta diferencia podría estar asociada a variaciones en el nivel educativo alcanzado por cada grupo. Dado que la educación constituye uno de los principales determinantes del ingreso en el modelo minceriano, resulta pertinente analizar su distribución por sexo a nivel nacional, lo cual se presenta en la Tabla 5.

Tabla 5. Nivel educativo por sexo (ENOE 2025-I, nivel nacional)

NIVEL EDUCATIVO	MUJERES (%)	HOMBRES (%)	TOTAL (%)
Preescolar	0.03%	0.05%	0.04%
Primaria	13.19%	17.04%	15.45%
Secundaria	25.52%	28.51%	27.27%
Preparatoria	23.14%	25.00%	24.23%
Carrera técnica	4.13%	2.02%	2.89%
Normal	0.16%	0.08%	0.11%
Profesional	28.55%	22.65%	25.10%
Maestría	2.69%	1.79%	2.16%
Doctorado	0.39%	0.31%	0.34%
No especificado	2.19%	2.55%	2.40%

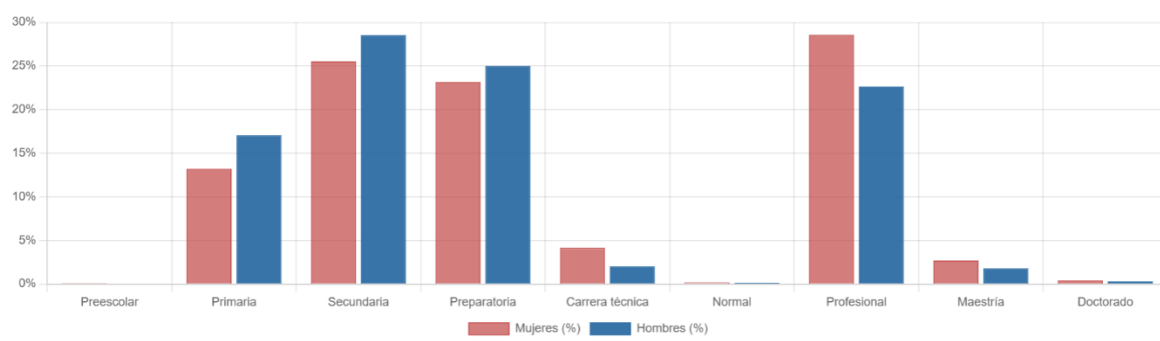
Fuente. Elaboración propia con base en datos de la ENOE 2025-I. La variable “no especificado” está construida con los datos agrupados de las variables sin dato, y “no sabe”.

La Tabla 5 presenta la distribución del nivel educativo por sexo a nivel nacional. Se observa que las mujeres muestran una mayor participación relativa en los niveles de educación superior, particularmente en el nivel profesional (29.07 % frente a 23.37 % en hombres) y en maestría (2.74 % frente a 1.85 %). En contraste, los hombres presentan una mayor concentración en los niveles de primaria, secundaria y preparatoria.

Estos resultados evidencian que, al menos en términos de escolaridad formal, las mujeres no se encuentran en desventaja respecto a los hombres; por el contrario, muestran una mayor presencia en los niveles educativos más altos. Por tanto, la brecha salarial identificada previamente no puede atribuirse exclusivamente a diferencias en el nivel educativo alcanzado.

Para facilitar la comparación visual entre ambos grupos, la Figura 3 representa gráficamente esta distribución, permitiendo apreciar de manera más clara las diferencias relativas en cada nivel educativo.

Figura 3. Distribución del nivel educativo por sexo (ENOE 2025-I, nivel nacional)



Fuente. Elaboración propia con base en datos de la ENOE 2025-I. La gráfica excluye la variable “No especificado”.

La representación gráfica confirma las diferencias descritas en la tabla, destacando la mayor participación femenina en educación superior y la mayor concentración masculina en niveles educativos básicos y medios.

Además del nivel educativo, la experiencia laboral constituye uno de los principales determinantes del ingreso en la teoría del capital humano (Mincer, 1974; Becker, 1964).

En este sentido, resulta pertinente analizar si existen diferencias en la experiencia laboral potencial entre hombres y mujeres, ya que esta variable permite aproximar la acumulación de trayectoria laboral de los individuos y su posible relación con las diferencias salariales observadas. Por ello, resulta pertinente analizar su distribución por sexo, lo cual se presenta

en la Tabla 6.

Tabla 6. *Experiencia laboral potencial por sexo (ENOE 2025-I, nivel nacional)*

SEXO	MEDIA	MEDIANA	DESVIACIÓN ESTÁNDAR	MÍNIMO	MÁXIMO	OBSERVACIONES
Mujeres	22	20	15	0	86	78,977
Hombres	22	21	16	0	86	111,526

Fuente. Elaboración propia con base en los datos de la ENOE 2025-I.

La Tabla 6 presenta los estadísticos descriptivos de la experiencia laboral potencial desagregados por sexo a nivel nacional. Se observa que la experiencia promedio es idéntica para hombres y mujeres, situándose en 22 años en ambos casos. Sin embargo, al analizar la mediana, se identifica una ligera diferencia, con un valor de 21 años para los hombres y 20 años para las mujeres.

En términos de dispersión, ambos grupos presentan niveles similares de variabilidad, con desviaciones estándar de 16 años para los hombres y 15 años para las mujeres, lo que indica una heterogeneidad comparable en las trayectorias laborales. Asimismo, los valores mínimo y máximo coinciden en ambos grupos, lo que sugiere que el rango de experiencia laboral potencial es prácticamente el mismo para hombres y mujeres a nivel nacional.

En este sentido, los resultados muestran que no existen diferencias relevantes en la experiencia laboral potencial entre hombres y mujeres, lo cual resulta particularmente significativo dado que esta variable constituye uno de los principales determinantes del ingreso según la teoría del capital humano. La ausencia de brechas sustanciales en esta dimensión sugiere que las diferencias salariales observadas previamente no pueden atribuirse a desigualdades en la experiencia laboral.

En conjunto, el análisis descriptivo realizado a partir de la ENOE a nivel nacional permite identificar la existencia de una brecha salarial entre hombres y mujeres, así como caracterizar las principales variables sociodemográficas y laborales asociadas al ingreso. Si bien los resultados muestran diferencias en la participación laboral y en los niveles educativos,

también evidencian que las mujeres presentan una presencia significativa en los niveles de educación superior. Asimismo, las diferencias en experiencia laboral potencial entre ambos grupos resultan relativamente moderadas.

Estos hallazgos sugieren que las disparidades salariales observadas no pueden atribuirse exclusivamente a diferencias en capital humano medido a través de la educación o la experiencia. En consecuencia, el análisis descriptivo, aunque necesario, resulta insuficiente para determinar qué proporción de la brecha salarial puede explicarse por características observables y cuál podría estar asociada a diferencias en los retornos de dichas características.

Por ello, en los siguientes apartados se procederá a aplicar el método de descomposición Oaxaca-Blinder, con el objetivo de cuantificar la parte explicada y no explicada de la brecha salarial, permitiendo así un análisis más profundo de las posibles fuentes de desigualdad en el mercado laboral mexicano.

3.5 Análisis de la muestra seleccionada: Banca y Seguros, edad de 23 a 27 años.

A partir de la base nacional depurada presentada en el apartado anterior, se seleccionaron 269 observaciones correspondientes a personas ocupadas en Banca y Seguros con edades comprendidas entre 23 y 27 años. Esta delimitación responde al objetivo de analizar la brecha salarial en un conjunto joven, reduciendo la heterogeneidad asociada al ciclo de vida laboral y a la diversidad sectorial observada a nivel agregado (Blau Y Kahn, 2017).

La submuestra está conformada por 140 mujeres (52.04 %) y 129 hombres (47.96 %), lo que refleja una distribución relativamente equilibrada entre ambos grupos. Esta composición permite realizar comparaciones salariales con menor sesgo derivado de diferencias en tamaño muestral, facilitando un análisis más preciso de las posibles disparidades de ingreso dentro del sector.

Al concentrarse en un grupo etario específico y en un sector formal del mercado laboral, se espera que variables como la experiencia laboral potencial y el nivel educativo presenten menor dispersión que a nivel nacional, lo cual permitirá evaluar con mayor claridad si la

brecha salarial persiste incluso bajo condiciones relativamente homogéneas.

Tabla 7. Composición por sexo (Banca y Seguros)

SEXO	OBSERVACIONES	PORCENTAJE
Mujeres	140	52%
Hombres	129	48%
Total	269	100%

Fuente. Elaboración propia con base en datos de la Encuesta Nacional de Ocupación y Empleo (ENOE 2025-I). La muestra incluye únicamente personas ocupadas en Banca y Seguros con edades entre 23 y 27 años, seleccionadas a partir de la base nacional depurada.

Con el objetivo de caracterizar las principales variables sociodemográficas y laborales de esta submuestra, la Tabla 8 presenta los estadísticos descriptivos correspondientes a Banca y Seguros para individuos de entre 23 y 27 años. Este análisis permite examinar el comportamiento general de variables como escolaridad, experiencia laboral potencial, salario mensual y horas trabajadas dentro de un grupo relativamente homogéneo en términos de edad y sector económico.

Tabla 8. Estadísticos descriptivos de la muestra seleccionada: Banca y Seguros (23–27 años)

VARIABLE	MEDIA	MEDIANA	DESV. ESTÁNDAR	MÍNIMO	MÁXIMO	OBSERVACIONES
Edad (años)	25	26	1	23	27	269
Escolaridad	14	16	2	6	18	269
Experiencia laboral en años (proxy)	5	5	3	0	15	269
Salario mensual (MXN)	13,753	12,546	5,864	8,364	50,184	201
Horas trabajadas por semana	44	44	9	0	60	268

Fuente. Elaboración propia con base en datos de la ENOE 2025-I. Los estadísticos descriptivos corresponden a la submuestra de personas ocupadas en la actividad bancaria y de seguros con edades entre 23 y 27 años. La variable de años de escolaridad fue construida mediante una equivalencia en años a partir del nivel educativo reportado en la encuesta. A partir de esta variable, la experiencia laboral potencial fue estimada como: $(edad - años\ de\ escolaridad - 6)$. Los valores no especificados fueron excluidos de los cálculos.

Después de haber caracterizado las principales variables sociodemográficas y laborales de la

submuestra, resulta pertinente analizar si existen diferencias salariales entre hombres y mujeres dentro de este grupo específico. Para ello, la Tabla 9 presenta los estadísticos descriptivos del salario mensual desagregados por sexo para trabajadores de Banca y Seguros con edades entre 23 y 27 años.

Tabla 9. Salario mensual por sexo (Banca y Sequros, 23–27 años)

SEXO	MEDIA (MXN)	MEDIANA (MXN)	DESV. ESTÁNDAR	MÍNIMO	MÁXIMO	OBSERVACIONES
● Mujeres	13,171	12,546	4,968	8,364	33,456	140
● Hombres	14,415	12,546	6,707	8,364	50,184	129
Brecha salarial 8.62%						

Fuente. Elaboración propia con base en datos de la ENOE 2025-I. Los estadísticos corresponden a trabajadores de Banca y Seguros con edades entre 23 y 27 años. La brecha salarial se calculó como la diferencia relativa del salario promedio femenino respecto al masculino. $((Media_{hombres} - Media_{mujeres} / Media_{hombres}) \times 100.)$

La Tabla 9 presenta los estadísticos descriptivos del salario mensual desagregados por sexo para la submuestra correspondiente a trabajadores del sector Banca y Seguros con edades entre 23 y 27 años. Se observa que el salario promedio de los hombres asciende a \$14,415 pesos, mientras que el de las mujeres es de \$13,171 pesos. Esta diferencia representa una brecha salarial de aproximadamente 8.62 %, lo que indica que, en promedio, las mujeres perciben ingresos menores que los hombres dentro de este sector.

Sin embargo, al comparar estos resultados con la brecha salarial observada a nivel nacional (13.84 %), se observa que la diferencia de ingresos entre hombres y mujeres dentro de Banca y Seguros es considerablemente menor. Esto sugiere que, en un entorno laboral más homogéneo y dentro de un sector formal, las disparidades salariales tienden a reducirse en comparación con el conjunto del mercado laboral.

Por otro lado, la mediana salarial es idéntica para ambos grupos (\$12,546 pesos), lo cual indica que en el centro de la distribución salarial no existen diferencias significativas entre hombres y mujeres. La diferencia observada en el salario promedio parece estar influenciada principalmente por la mayor dispersión salarial entre los hombres, quienes presentan una

desviación estándar más alta y un valor máximo superior.

Estos resultados sugieren que, aunque la brecha salarial persiste dentro del sector analizado, su magnitud es menor y podría estar asociada a diferencias en los niveles superiores de ingreso más que a una desigualdad generalizada en toda la distribución salarial.

Con el fin de profundizar en el análisis de las diferencias salariales observadas, resulta necesario examinar si hombres y mujeres dentro de esta submuestra presentan características productivas similares. En la literatura de economía laboral, variables asociadas al capital humano, como el nivel educativo, suelen desempeñar un papel importante en la determinación de los ingresos (Mincer, 1974; Becker, 1964). Por ello, analizar la distribución del nivel de escolaridad entre ambos grupos permite evaluar si las diferencias salariales pueden estar relacionadas con disparidades en la formación académica de los trabajadores.

En este contexto, la Tabla 10 presenta la distribución del nivel educativo por sexo para los trabajadores de Banca y Seguros con edades entre 23 y 27 años.

Tabla 10. Nivel educativo por sexo (Banca y Seguros, 23–27 años)

NIVEL EDUCATIVO	MUJERES (%)	HOMBRES (%)
Preescolar	0.00	0.00
Primaria	0.71	0.78
Secundaria	7.14	4.65
Preparatoria	25.00	27.91
Carrera técnica	1.43	2.33
Normal	0.00	0.00
Profesional	65.00	64.34
Maestría	0.71	0.00
Doctorado	0.00	0.00

Fuente. Elaboración propia con base en datos de la ENOE 2025-I. Los estadísticos corresponden a trabajadores de Banca y Seguros con edades entre 23 y 27 años.

La Tabla 10 presenta la distribución del nivel educativo por sexo para los trabajadores de Banca y Seguros con edades entre 23 y 27 años. En términos generales, se observa una alta concentración de ambos grupos en niveles educativos medio superior y superior, lo cual es

consistente con las características del sector analizado, que suele requerir personal con mayor formación académica.

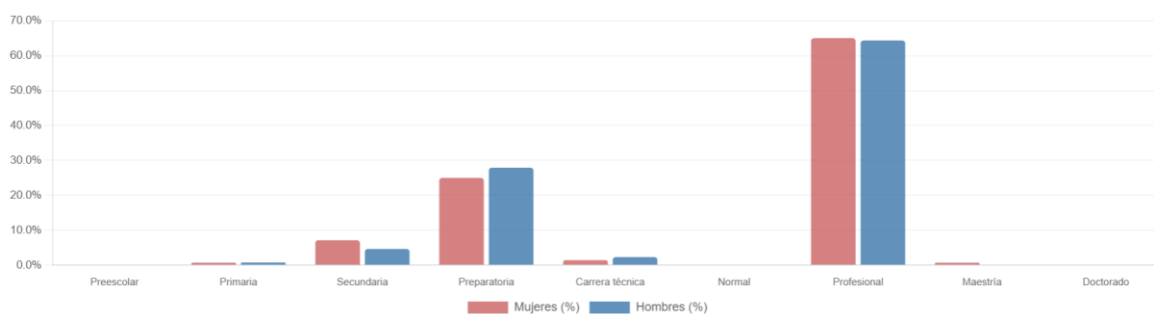
En particular, el nivel profesional concentra la mayor proporción de trabajadores en ambos casos, representando el 65.00 % de las mujeres y el 64.34 % de los hombres. Asimismo, el nivel medio superior (preparatoria) constituye el segundo nivel educativo más frecuente, con 25.00 % de las mujeres y 27.91 % de los hombres. Los niveles educativos inferiores presentan una participación considerablemente menor dentro de la muestra, mientras que los estudios de posgrado aparecen únicamente de forma marginal.

En conjunto, estos resultados muestran que la distribución del nivel educativo entre hombres y mujeres dentro de la submuestra es altamente similar. Esta evidencia sugiere que las diferencias salariales observadas previamente no parecen estar asociadas a disparidades significativas en la formación académica de los trabajadores, lo cual refuerza la necesidad de analizar otros factores relacionados con la determinación de los ingresos (Blau Y Kahn, 2017).

Con el fin de visualizar de manera más clara la distribución del nivel educativo entre hombres y mujeres dentro de la submuestra analizada, la Figura 4. presenta una representación gráfica de los porcentajes correspondientes a cada nivel educativo. La utilización de esta representación permite comparar de forma más intuitiva la concentración de trabajadores en los distintos niveles de formación académica y observar con mayor claridad las similitudes existentes entre ambos grupos.

Como puede apreciarse en la figura, la mayor concentración de trabajadores se ubica en el nivel profesional, seguido por preparatoria, mientras que los niveles educativos inferiores y de posgrado presentan una participación considerablemente menor dentro de la muestra. En términos generales, la gráfica confirma que la distribución educativa entre hombres y mujeres es altamente similar, lo que refuerza la evidencia presentada en la tabla anterior respecto a la ausencia de diferencias significativas en el capital humano educativo entre ambos grupos.

Figura 4. Distribución del nivel educativo por sexo en Banca y Seguros (23-27 años)



Fuente. Elaboración propia con base en datos de la ENOE 2025-I.

La Figura 4 es una representación gráfica de la distribución del nivel educativo por sexo dentro de la submuestra analizada. La visualización permite apreciar de forma clara la concentración de trabajadores en los niveles de preparatoria y, principalmente, profesional, así como la baja participación de los niveles educativos inferiores y de posgrado dentro del sector analizado.

En conjunto, los resultados presentados en la Tabla 10 y en la Figura 4 muestran que la distribución del nivel educativo entre hombres y mujeres dentro de la submuestra analizada es altamente similar. Dado que la formación académica constituye uno de los principales componentes del capital humano en la determinación de los ingresos laborales, estos resultados sugieren que las diferencias salariales observadas previamente difícilmente pueden explicarse únicamente por disparidades en el nivel educativo de los trabajadores.

No obstante, además de la educación, la literatura en economía laboral reconoce a la experiencia laboral como otro factor fundamental en la determinación de los salarios (Mincer, 1974). En este sentido, resulta pertinente examinar si existen diferencias en la experiencia acumulada entre hombres y mujeres dentro de la submuestra analizada. Para ello, la Tabla 11 presenta los estadísticos descriptivos de la experiencia laboral potencial desagregados por sexo para los trabajadores de Banca y Seguros con edades entre 23 y 27 años.

Tabla 11. *Experiencia laboral potencial por sexo*

SEXO	MEDIA (AÑOS)	MEDIANA (AÑOS)	DESV. ESTÁNDAR	MÍNIMO	MÁXIMO	OBSERVACIONES
● Mujeres	5	4	3	0	15	140
● Hombres	5	5	3	1	15	129

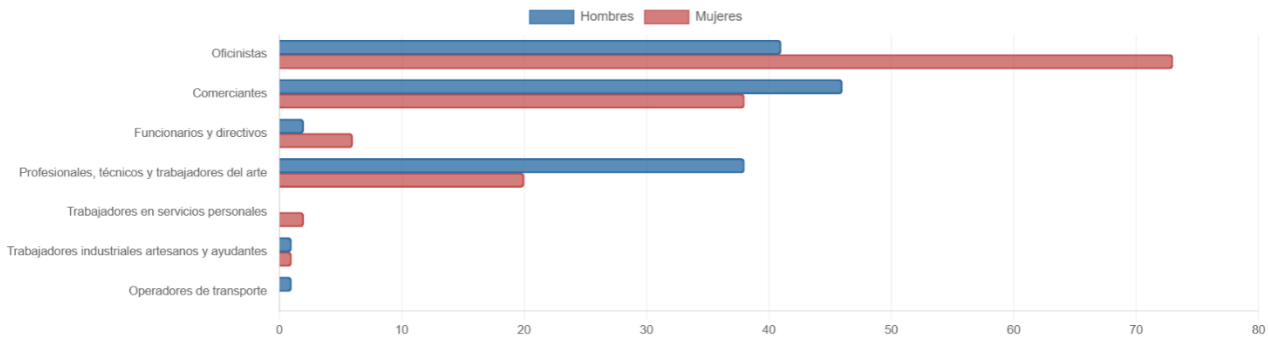
Fuente. Elaboración propia con base en datos de la ENOE 2025-I. La experiencia laboral (proxy) se estimó como edad menos años de escolaridad menos seis.

La Tabla 11 presenta los estadísticos descriptivos de la experiencia laboral potencial desagregados por sexo para la submuestra correspondiente a trabajadores de Banca y Seguros con edades entre 23 y 27 años. Los resultados muestran que tanto hombres como mujeres presentan un nivel promedio de experiencia laboral potencial similar, con una media de 5 años en ambos casos.

Asimismo, la mediana de la experiencia laboral es de 4 años para las mujeres y 5 años para los hombres, mientras que la desviación estándar es idéntica para ambos grupos (3 años), lo que indica una dispersión comparable en la distribución de esta variable. Estos resultados son consistentes con las características de la cohorte analizada, ya que al tratarse de trabajadores jóvenes que se encuentran en las primeras etapas de su trayectoria profesional, la acumulación de experiencia laboral tiende a ser relativamente limitada y homogénea.

En conjunto, la evidencia presentada sugiere que las diferencias salariales observadas previamente entre hombres y mujeres dentro del sector analizado difícilmente pueden atribuirse a disparidades significativas en la experiencia laboral potencial de los trabajadores.

Figura 5. *Distribución de la muestra por ocupación y sexo (Banca y Seguros, 23-27 años)*



Fuente. Elaboración propia con base en datos de la ENOE 2025-I.

Con el objetivo de complementar el análisis de las características laborales de la submuestra, la Figura 5 presenta la distribución de los trabajadores por tipo de ocupación y género. Esta representación permite identificar posibles diferencias en la estructura ocupacional que no son capturadas por las variables tradicionales de capital humano.

Como se observa en la figura, existe una concentración importante de mujeres en ocupaciones administrativas, particularmente en la categoría de oficinistas, mientras que los hombres presentan una mayor participación en ocupaciones como comerciantes y profesionales, técnicos y trabajadores del arte.

En el caso de los puestos de mayor jerarquía, como funcionarios y directivos, se observa una ligera mayor presencia femenina en términos absolutos; sin embargo, esta categoría representa una proporción muy reducida dentro de la muestra total, por lo que su impacto en la estructura ocupacional general es limitado.

En conjunto, estos resultados sugieren la existencia de una segmentación ocupacional por género dentro del Banca y Seguros. En este sentido, aun cuando hombres y mujeres presentan niveles similares de educación y experiencia laboral, su concentración en distintos tipos de ocupaciones puede influir en las diferencias salariales observadas (Anker, 1997; Blau Y Kahn, 2017).

En resumen, el análisis descriptivo de la submuestra correspondiente a trabajadores de Banca y Seguros con edades entre 23 y 27 años muestra que, si bien existe una brecha salarial entre hombres y mujeres, las principales variables asociadas al capital humano (como son el nivel educativo y la experiencia laboral potencial) presentan distribuciones muy similares entre ambos grupos. Estos resultados sugieren que las diferencias salariales observadas no pueden explicarse únicamente por variaciones en las características productivas de los individuos.

En este contexto, resulta pertinente aplicar una metodología que permita descomponer la brecha salarial y distinguir qué proporción de esta puede atribuirse a diferencias en las características observables de los trabajadores y qué parte permanece sin explicación por dichos factores. Con este propósito, en el siguiente apartado se implementa el método de

descomposición Oaxaca-Blinder, el cual permite analizar de manera más detallada los determinantes de la brecha salarial entre hombres y mujeres dentro del sector estudiado (Oaxaca, 1973; Blinder, 1973).

Los resultados descriptivos evidencian la existencia de diferencias salariales entre hombres y mujeres, así como ciertas particularidades en variables relacionadas con la educación, la experiencia laboral y las condiciones de empleo. Sin embargo, el análisis descriptivo por sí solo no permite determinar qué proporción de dichas diferencias puede explicarse por características observables ni qué parte permanece sin explicación. Por ello, en el siguiente capítulo se aplicará la metodología Oaxaca-Blinder con el propósito de profundizar en el estudio de la brecha salarial de género dentro de Banca y Seguros.

CAPÍTULO 4. APLICACIÓN DE LA TÉCNICA ECONOMETRICA OAXACA-BLINDER PARA EL CASO DE ESTUDIO.

Con base en la información presentada en los capítulos anteriores, este capítulo expone los resultados obtenidos mediante la aplicación de la metodología Oaxaca-Blinder. En primer lugar, se presentan las especificaciones de los modelos econométricos considerados en el análisis y posteriormente se examinan los resultados de la descomposición de la brecha salarial. El objetivo es identificar qué proporción de las diferencias salariales observadas puede atribuirse a características asociadas al capital humano y qué proporción permanece sin explicar por los factores incluidos en el modelo.

4.1 Selección de variables

Para la presente investigación, la variable dependiente corresponde al salario mensual del individuo, expresado en pesos mexicanos. Esta variable se selecciona debido a que refleja de manera directa el nivel de ingreso que perciben las personas por su trabajo dentro del de Banca y Seguros, lo que permite analizar las diferencias salariales entre grupos y los factores que las explican. A partir de esta variable se busca identificar cómo influyen distintos aspectos personales, laborales y regionales en la remuneración que recibe cada individuo.

En cuanto a las variables independientes, su selección se realiza con base en la literatura sobre el mercado laboral y la teoría del capital humano, la cual establece que factores como la educación, la experiencia y otras características individuales influyen en la productividad y, por tanto, en el salario. Asimismo, se incorpora una variable que permite capturar la intensidad del trabajo, con el fin de complementar el análisis desde una perspectiva más integral.

Las variables incluidas en el modelo son las siguientes:

- **Nivel de escolaridad potencial (proxy):** estimada mediante una equivalencia en años de estudio a partir del nivel educativo reportado por los individuos en la Encuesta Nacional de Ocupación y Empleo (ENOE). Dado que la encuesta no

proporciona directamente una variable numérica continua que represente los años efectivos de escolaridad de los trabajadores, se construyó una variable proxy asignando a cada nivel educativo los años de estudio correspondientes dentro del sistema educativo mexicano. De esta forma, se consideraron 6 años para primaria completa, 9 años para secundaria completa, 12 años para educación media superior, 16 años para licenciatura, 18 años para maestría y 21 años para doctorado (La equivalencia completa utilizada para la construcción de esta variable se presenta en el Anexo A). Esta aproximación permite representar de manera indirecta el nivel de formación académica y la acumulación de capital humano de los individuos, constituyendo uno de los principales factores utilizados en los modelos salariales para explicar las diferencias en los ingresos laborales.

- **Experiencia laboral potencial (proxy):** estimada en años como una aproximación de la experiencia laboral acumulada por el individuo. Dado que la Encuesta Nacional de Ocupación y Empleo no proporciona directamente una variable que mida la trayectoria laboral completa de los trabajadores, esta se aproxima mediante una medida de experiencia laboral potencial, calculada a partir de la edad y los años de escolaridad. Esta variable busca capturar el tiempo que un individuo podría haber permanecido en el mercado laboral después de finalizar su formación académica y constituye un componente central en los modelos de capital humano utilizados para explicar las diferencias en los ingresos laborales.
- **Promedio de horas trabajadas por semana:** variable que refleja la intensidad laboral del individuo. Su inclusión permite capturar diferencias en el salario asociadas al tiempo dedicado al trabajo, ya que, en general, una mayor cantidad de horas trabajadas puede traducirse en mayores niveles de ingreso.
- **Sexo:** variable categórica codificada de forma binaria (0 = hombre, 1 = mujer), incorporada con el objetivo de analizar la existencia de diferencias salariales atribuibles al sexo, aun cuando hombres y mujeres cuenten con características laborales similares. Esta variable es fundamental para estudiar la brecha salarial de género dentro del sector.

Es importante señalar que, si bien la edad es una variable comúnmente utilizada en modelos de ingresos, en el presente estudio no se incluye de manera explícita dentro de la especificación econométrica. Esto se debe a que: el análisis se centra en una submuestra de individuos con edades comprendidas entre 23 y 27 años, lo cual reduce significativamente la variabilidad de esta variable. Asimismo, la variable edad no fue incorporada de manera independiente dentro de las especificaciones finales del modelo. Esto se debe a que la experiencia laboral potencial se construyó a partir de la edad y los años de escolaridad, por lo que ambas variables presentan una relación directa que podría generar problemas de redundancia o multicolinealidad en la estimación. Adicionalmente, la muestra analizada se restringe a individuos de entre 23 y 27 años de edad, lo que reduce considerablemente la variabilidad de esta característica dentro de la población de estudio. Por estas razones, se optó por utilizar únicamente la experiencia laboral potencial como aproximación a la trayectoria laboral de los individuos.

En conjunto, estas variables permiten analizar el comportamiento del salario desde una perspectiva integral, considerando tanto los factores asociados al capital humano como la intensidad del trabajo. Su selección busca aportar una comprensión más completa de las desigualdades salariales y de los elementos que contribuyen a la brecha entre hombres y mujeres en Banca y Seguros.

4.2 Propuesta de modelo

La presente investigación propone la utilización del método de descomposición Oaxaca-Blinder (1973) como herramienta central para el análisis de la brecha salarial entre hombres y mujeres dentro de Banca y Seguros en México, a partir de la información obtenida de la Encuesta Nacional de Ocupación y Empleo (ENOE).

El método Oaxaca-Blinder es ampliamente reconocido en la literatura económica y social por su capacidad para identificar las fuentes que explican las desigualdades salariales. A diferencia de otros modelos de regresión que solo estiman el efecto de las variables sobre el salario, esta técnica permite dividir la diferencia promedio entre los ingresos de dos grupos (en este caso, hombres y mujeres) en dos componentes principales: la parte explicada,

asociada a las diferencias observables en las características de los trabajadores, como la educación, la experiencia laboral o la intensidad del trabajo; y la parte no explicada, que refleja diferencias en los retornos a esas características y puede interpretarse como desigualdad estructural o la presencia de factores no observables.

El modelo se sustenta en la teoría del capital humano, desarrollada por Jacob Mincer (1974), la cual establece que el salario depende del nivel educativo y de la experiencia laboral acumulada. Bajo esta perspectiva, los trabajadores con mayor formación o experiencia tienden a recibir salarios más altos, ya que son considerados más productivos. No obstante, diversos estudios han demostrado que, incluso con características similares, las mujeres perciben menores ingresos que los hombres, lo que sugiere la presencia de factores no observables o mecanismos de desigualdad en el mercado laboral.

A partir de este enfoque, el modelo Oaxaca-Blinder combina la ecuación minceriana del salario con una descomposición que permite cuantificar cuánto de la brecha salarial puede justificarse por diferencias en el capital humano y cuánto se asocia a diferencias en los rendimientos que el mercado otorga a esas características. En el presente estudio, la ecuación salarial se especifica de la siguiente manera:

$$\ln(\text{Salario}_i) = \beta_0 + \beta_1 \text{Escolaridad}_i + \beta_2 \text{Experiencia}_i + \beta_3 \text{Horas}_i + \varepsilon_i$$

En esta ecuación, el logaritmo del salario mensual se expresa en función del nivel educativo, la experiencia laboral potencial y el promedio de horas trabajadas por semana. El uso del logaritmo natural del salario permite modelar de manera más adecuada la relación entre los ingresos y sus determinantes, dado que los salarios suelen presentar distribuciones asimétricas. Asimismo, esta transformación facilita la interpretación de los coeficientes estimados, los cuales pueden entenderse como variaciones porcentuales aproximadas en el salario ante cambios en las variables explicativas.

Desde el punto de vista económico, el coeficiente asociado a la escolaridad refleja el incremento porcentual en el salario derivado de un año adicional de estudio, manteniendo constantes las demás variables. Por su parte, la experiencia laboral potencial captura el efecto de la trayectoria laboral sobre los ingresos, mientras que la variable de horas trabajadas

permite incorporar la intensidad del trabajo, reconociendo que una mayor dedicación en términos de tiempo puede traducirse en mayores niveles de ingreso.

Con base en esta especificación, se estiman dos regresiones separadas, una para hombres y otra para mujeres, con el objetivo de obtener coeficientes específicos para cada grupo. Esto permite analizar no solo las diferencias en las características promedio de los individuos, sino también en la forma en que el mercado remunera dichas características.

Formalmente, las ecuaciones salariales para cada grupo pueden expresarse de la siguiente manera:

$$\begin{aligned}\ln(\text{Salario}_h) &= X_h\beta_h + \varepsilon_h \\ \ln(\text{Salario}_m) &= X_m\beta_m + \varepsilon_m\end{aligned}$$

donde X_h y X_m representan las matrices de variables explicativas para hombres y mujeres, respectivamente, mientras que β_h y β_m corresponden a los vectores de coeficientes estimados para cada grupo. Esta desagregación permite identificar si existen diferencias en los retornos a las características observables entre hombres y mujeres, lo cual constituye el elemento central para la posterior aplicación de la descomposición Oaxaca-Blinder.

Una vez estimadas las ecuaciones salariales, se aplica la descomposición Oaxaca-Blinder, que se expresa de la siguiente manera:

$$\ln(\bar{\text{Salario}}_h) - \ln(\bar{\text{Salario}}_m) = (\bar{X}_h - \bar{X}_m)\hat{\beta}_h + \bar{X}_m(\hat{\beta}_h - \hat{\beta}_m)$$

En esta expresión, el primer término representa la parte explicada de la brecha salarial, atribuida a diferencias observables entre los individuos, como la escolaridad, la experiencia o las horas trabajadas. El segundo término corresponde a la parte no explicada, asociada a diferencias en los coeficientes, lo que puede interpretarse como evidencia de desigualdad en los retornos del mercado laboral o la presencia de factores no observados.

El modelo es estimado mediante el método de Mínimos Cuadrados Ordinarios (MCO), y la descomposición se lleva a cabo a partir de las estimaciones obtenidas para cada grupo. En primer lugar, se estima la ecuación salarial para el total de la muestra con el fin de observar la relación general entre las variables de estudio y el salario. Posteriormente, se realiza estimaciones por separado para hombres y mujeres, lo que permitirá comparar los coeficientes y evaluar las diferencias en los rendimientos de las características laborales. A partir de estas estimaciones, se calcula la brecha promedio de salarios y se aplica la descomposición Oaxaca-Blinder para separar dicha brecha en sus componentes explicables y no explicables.

Con base en la teoría y la evidencia empírica, se espera que las variables nivel de escolaridad, experiencia laboral y horas trabajadas tengan un efecto positivo sobre el salario. Asimismo, se anticipa que las mujeres presenten, en promedio, salarios más bajos que los hombres, incluso cuando comparten características observables similares. La parte no explicada de la descomposición Oaxaca-Blinder reflejaría así la existencia de una brecha salarial por razones de género.

El modelo Oaxaca-Blinder permite identificar de manera cuantitativa los factores asociados a las diferencias salariales y estimar el componente atribuible a desigualdad. No obstante, sus resultados dependen de la calidad y disponibilidad de la información empleada. Además, al basarse en variables observables, el modelo no captura completamente otros factores estructurales o sociales que también pueden influir en la desigualdad salarial.

A pesar de estas limitaciones, la aplicación de esta metodología en el contexto mexicano representa una herramienta sólida para analizar el comportamiento de la brecha salarial en Banca y Seguros, contribuyendo al diseño de políticas laborales orientadas a la equidad y a la igualdad de oportunidades.

No obstante, antes de proceder al análisis de resultados, resulta necesario evaluar si la especificación econométrica adoptada cumple con los supuestos que garantizan la validez de las estimaciones obtenidas. En este sentido, la correcta interpretación de los coeficientes y de la descomposición Oaxaca-Blinder depende de que el modelo presente un comportamiento

estadísticamente adecuado.

Por ello, en el siguiente apartado se presentan las pruebas de diagnóstico aplicadas al modelo estimado, con el objetivo de verificar la consistencia de los resultados y asegurar la solidez del análisis econométrico.

4.3 Pruebas de diagnóstico

Una vez estimado el modelo econométrico propuesto, resulta fundamental evaluar si se cumplen los supuestos que garantizan la validez de las estimaciones obtenidas. En este sentido, las pruebas de diagnóstico constituyen una etapa esencial dentro del análisis econométrico, ya que permiten identificar posibles problemas en la especificación del modelo y verificar la confiabilidad de los resultados.

En modelos estimados mediante el método de Mínimos Cuadrados Ordinarios (MCO), el cumplimiento de supuestos como la homocedasticidad, la normalidad de los errores y la ausencia de multicolinealidad resulta clave para asegurar la validez de la inferencia estadística (Gujarati y Porter, 2009). Cuando estos supuestos no se cumplen, los coeficientes estimados pueden seguir siendo insesgados, pero la eficiencia de los estimadores y la precisión de las pruebas de hipótesis pueden verse comprometidas.

Dado que en el presente estudio se estiman ecuaciones salariales de manera separada para hombres y mujeres, las pruebas de diagnóstico se aplican de forma diferenciada para cada grupo.

4.3.1 Heterocedasticidad: prueba de Breusch-Pagan

La heterocedasticidad se refiere a la presencia de varianza no constante en el término de error. Para evaluar este supuesto, se empleó la prueba de Breusch-Pagan, cuya hipótesis nula establece que la varianza de los residuos es constante (Breusch Y Pagan, 1979).

En el caso de los hombres, se aplica la prueba de Breusch-Pagan con el objetivo de evaluar

la estabilidad de la varianza de los errores en la ecuación salarial estimada.

Figura 6. Resultados de la prueba Breusch-Pagan para hombres



Fuente. Elaboración propia mediante procesamiento de datos en Python y visualización en entorno web.

Los resultados muestran un estadístico LM de 14.2190 con un valor p de 0.000817, así como un estadístico F de 5.3545 con un valor p de 0.000325. Dado que ambos valores p son menores al nivel de significancia del 5 %, se rechaza la hipótesis nula de homocedasticidad, lo que indica la presencia de heterocedasticidad en el modelo.

Este resultado sugiere que la varianza de los residuos no es constante, lo cual puede afectar la eficiencia de los estimadores y la validez de los errores estándar. No obstante, es importante señalar que los coeficientes estimados mediante MCO permanecen insesgados, aunque se recomienda el uso de errores estándar robustos para una inferencia más precisa.

En el caso de las mujeres, se aplica la prueba de Breusch-Pagan con el objetivo de evaluar la estabilidad de la varianza de los errores en la ecuación salarial estimada.

Figura 7. Resultados de la prueba Breush-Pagan para mujeres



Fuente. Elaboración propia mediante procesamiento de datos en Python y visualización en entorno web.

Los resultados muestran un estadístico LM de 2.7489 con un valor p de 0.252978, así como un estadístico F de 0.9053 con un valor p de 0.257186. Dado que ambos valores p son mayores al nivel de significancia del 5 %, no se rechaza la hipótesis nula de homocedasticidad, lo que indica que no existe evidencia de heterocedasticidad en el modelo. Este resultado sugiere que la varianza de los residuos puede considerarse constante, por lo que no se identifican problemas relevantes en términos de heterocedasticidad para este grupo. En consecuencia, los estimadores obtenidos mediante MCO pueden considerarse eficientes bajo este supuesto.

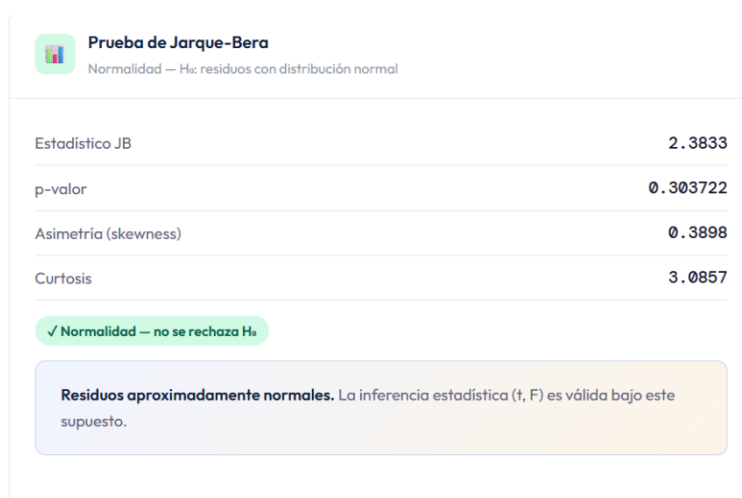
En conjunto, los resultados de la prueba de Breusch-Pagan muestran un comportamiento diferenciado entre hombres y mujeres. Mientras que en el caso de los hombres se detecta la presencia de heterocedasticidad, en el caso de las mujeres no se encuentra evidencia de este problema. Esta diferencia sugiere que la variabilidad de los errores no se distribuye de la misma forma entre ambos grupos, lo cual resulta relevante para la interpretación de las estimaciones y refuerza la importancia de analizar de manera separada las ecuaciones salariales por sexo.

4.3.2 Normalidad de los residuos: prueba de Jarque-Bera

La normalidad de los residuos es un supuesto relevante para la validez de las pruebas de hipótesis en modelos econométricos. Para evaluarlo, se empleó la prueba de Jarque-Bera, la cual permite analizar si los residuos del modelo siguen una distribución normal a partir de su asimetría y curtosis (Jarque Y Bera, 1987). La hipótesis nula establece que los residuos se distribuyen normalmente.

En el caso de los hombres, se aplica la prueba de Jarque-Bera con el objetivo de evaluar la distribución de los errores en la ecuación salarial estimada.

Figura 8. Resultados de la prueba Jarque-Bera para hombres



Fuente. Elaboración propia mediante procesamiento de datos en Python y visualización en entorno web.

Los resultados muestran un estadístico Jarque-Bera de 2.3833 con un valor p de 0.3037. Dado que el valor p es mayor al nivel de significancia del 5 %, no se rechaza la hipótesis nula de normalidad, lo que indica que los residuos pueden considerarse aproximadamente normales. Adicionalmente, el coeficiente de asimetría (skewness) presenta un valor de 0.3898, lo que sugiere una ligera asimetría positiva en la distribución de los residuos. Por su parte, la curtosis es de 3.0857, valor cercano a 3, lo cual indica que la distribución es similar a una normal en términos de concentración de datos alrededor de la media.

En conjunto, estos resultados sugieren que el modelo cumple razonablemente con el supuesto

de normalidad en este grupo, lo cual favorece la validez de la inferencia estadística, en particular en lo que respecta a la interpretación de los coeficientes y las pruebas de significancia.

En el caso de las mujeres, se aplica la prueba de Jarque-Bera con el objetivo de evaluar si los residuos de la ecuación salarial estimada siguen una distribución normal.

Figura 9. Resultados de la prueba Jarque-Bera para mujeres



Fuente. Elaboración propia mediante procesamiento de datos en Python y visualización en entorno web.

Los resultados muestran un estadístico Jarque-Bera de 15.2946 con un valor p de 0.000477. Dado que el valor p es menor al nivel de significancia del 5 %, se rechaza la hipótesis nula de normalidad, lo que indica que los residuos no siguen una distribución normal.

Adicionalmente, el coeficiente de asimetría (skewness) presenta un valor de 0.8049, lo que sugiere una asimetría positiva más marcada en la distribución de los residuos. Por su parte, la curtosis es de 3.9160, valor superior a 3, lo cual indica una mayor concentración de datos en torno a la media y colas más pesadas en comparación con una distribución normal.

Estos resultados sugieren que el modelo no cumple con el supuesto de normalidad en este grupo, lo cual puede afectar la precisión de la inferencia estadística. No obstante, considerando el tamaño de la muestra y las propiedades del estimador MCO, este

incumplimiento no invalida el modelo, aunque sí sugiere que los resultados deben interpretarse con cautela.

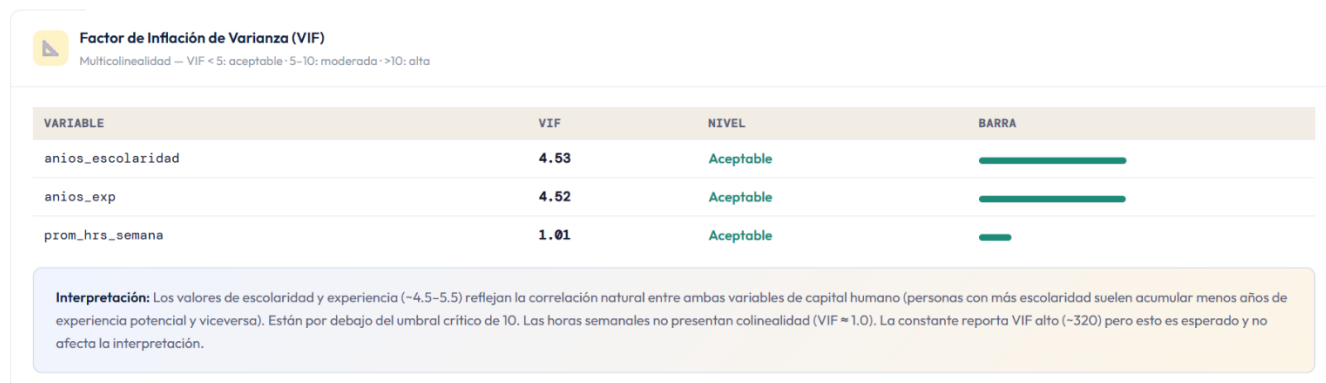
En conjunto, los resultados de la prueba de Jarque-Bera evidencian un comportamiento diferenciado entre hombres y mujeres. Mientras que en el caso de los hombres los residuos se aproximan a una distribución normal, en el caso de las mujeres se observan desviaciones significativas respecto a este supuesto, caracterizadas por una mayor asimetría y curtosis. Esta diferencia sugiere que la distribución de los errores no es homogénea entre ambos grupos, lo cual refuerza la importancia de analizar las ecuaciones salariales de manera separada por sexo. Asimismo, estos hallazgos son consistentes con el comportamiento observado en el análisis gráfico de residuos, lo que aporta evidencia adicional sobre la estructura del modelo estimado.

4.3.3 Multicolinealidad: Factor de Inflación de la Varianza (VIF)

La multicolinealidad se refiere a la existencia de una alta correlación entre las variables explicativas del modelo, lo cual puede afectar la estabilidad y precisión de los coeficientes estimados. Para evaluar este supuesto, se empleó el Factor de Inflación de la Varianza (VIF), el cual mide el grado en que la varianza de un coeficiente se incrementa debido a la colinealidad entre las variables independientes (Gujarati y Porter, 2009). En general, valores de VIF inferiores a 5 se consideran aceptables, entre 5 y 10 indican colinealidad moderada, y superiores a 10 representan un problema severo.

En el caso de los hombres, se calculó el VIF para cada una de las variables explicativas del modelo.

Figura 10. Resultados de la prueba Factor de Inflación de Varianza (VIF) para hombres



Fuente. Elaboración propia mediante procesamiento de datos en Python y visualización en entorno web.

Los resultados muestran valores de 4.53 para la escolaridad, 4.52 para la experiencia laboral y 1.01 para las horas trabajadas.

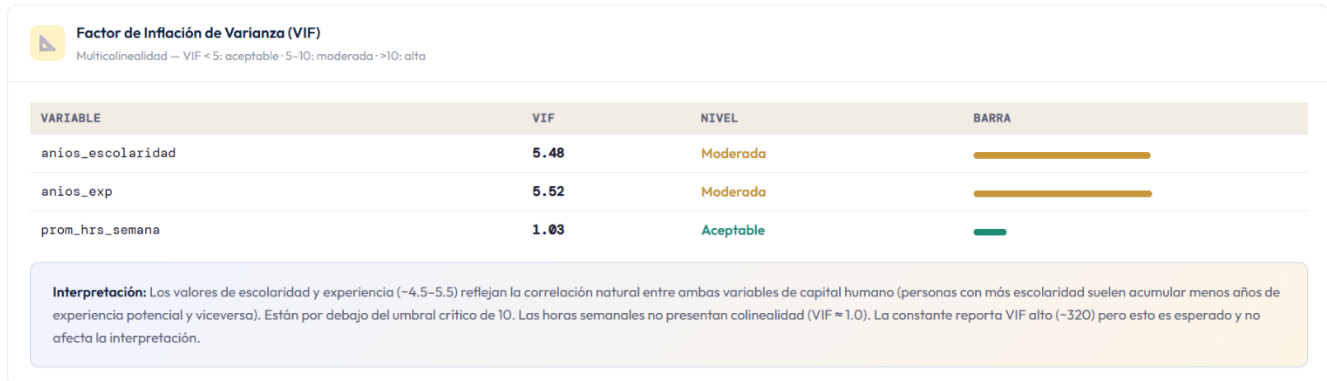
Dado que todos los valores se encuentran por debajo del umbral crítico de 5, no se identifican problemas relevantes de multicolinealidad en el modelo. En particular, la variable de horas trabajadas presenta un valor cercano a 1, lo que indica ausencia de correlación con las demás variables.

Asimismo, la correlación moderada observada entre la escolaridad y la experiencia es consistente con la teoría del capital humano, ya que ambas variables capturan dimensiones relacionadas del perfil laboral de los individuos. En este sentido, cierto grado de colinealidad entre ellas es esperado y no necesariamente constituye un problema, siempre que los niveles se mantengan dentro de rangos aceptables, como ocurre en el presente caso.

En consecuencia, se considera que el modelo para este grupo cumple adecuadamente con el supuesto de ausencia de multicolinealidad.

En el caso de las mujeres, se aplica el mismo análisis para evaluar la posible existencia de colinealidad entre las variables explicativas.

Figura 11. Resultados de la prueba Factor de Inflación de Varianza (VIF) para mujeres



Fuente. Elaboración propia mediante procesamiento de datos en Python y visualización en entorno web.

Los resultados muestran valores de 5.48 para la escolaridad, 5.52 para la experiencia laboral y 1.03 para las horas trabajadas.

En este caso, las variables de escolaridad y experiencia presentan valores ligeramente superiores a 5, lo que indica la presencia de una colinealidad moderada. Sin embargo, estos valores se mantienen por debajo del umbral crítico de 10, por lo que no representan un problema severo para la estimación del modelo.

Por su parte, la variable de horas trabajadas continúa mostrando un valor cercano a 1, lo que confirma que no presenta correlación significativa con las demás variables.

En conjunto, estos resultados sugieren que, si bien existe cierta relación entre las variables de capital humano, esta no compromete de manera significativa la estabilidad del modelo.

En conjunto, los resultados del análisis de multicolinealidad muestran que el modelo presenta un comportamiento adecuado en ambos grupos. Mientras que en el caso de los hombres no se identifican problemas relevantes de colinealidad, en el caso de las mujeres se observa una colinealidad moderada entre las variables de escolaridad y experiencia, lo cual es consistente con la relación teórica entre ambas.

No obstante, en ninguno de los casos se alcanzan niveles críticos que comprometan la validez

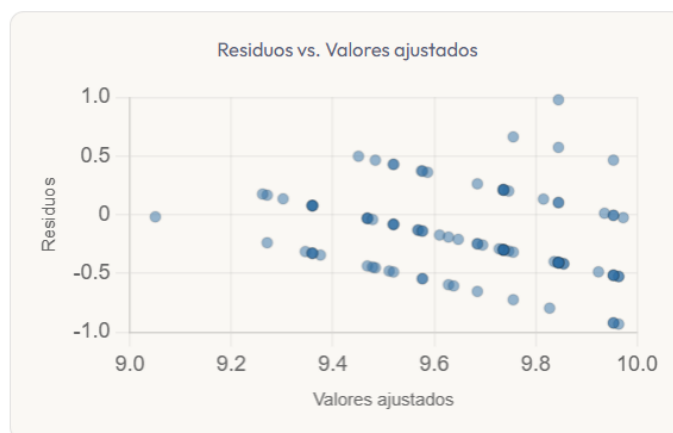
del modelo, por lo que se concluye que las variables incluidas pueden ser utilizadas de manera conjunta sin generar distorsiones significativas en la estimación de los coeficientes.

4.3.4 Análisis de residuos

El análisis de residuos constituye una herramienta fundamental para evaluar la adecuación del modelo econométrico, ya que permite identificar posibles problemas de especificación, patrones no capturados y desviaciones respecto a los supuestos clásicos. En este sentido, se analizaron tres representaciones gráficas: residuos contra valores ajustados, histogramas de residuos y gráficos Q-Q, con el objetivo de evaluar la homocedasticidad, la normalidad y el comportamiento general del modelo.

En el caso de los hombres, se realiza un análisis gráfico de los residuos con el fin de evaluar la adecuación del modelo estimado.

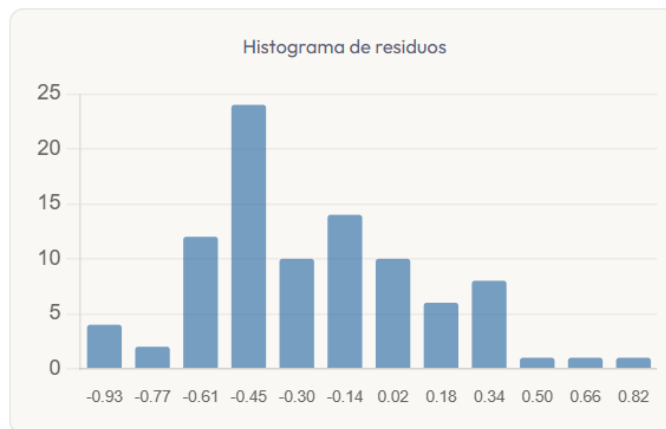
Figura 12. Gráfico de residuos vs. valores ajustados (hombres)



Fuente. Elaboración propia mediante procesamiento de datos en Python y visualización en entorno web.

En la figura 12, el gráfico de residuos contra valores ajustados, se observa una ligera dispersión creciente a medida que aumentan los valores ajustados, lo cual sugiere la posible presencia de heterocedasticidad. Este resultado es consistente con lo observado previamente en la prueba de Breusch-Pagan.

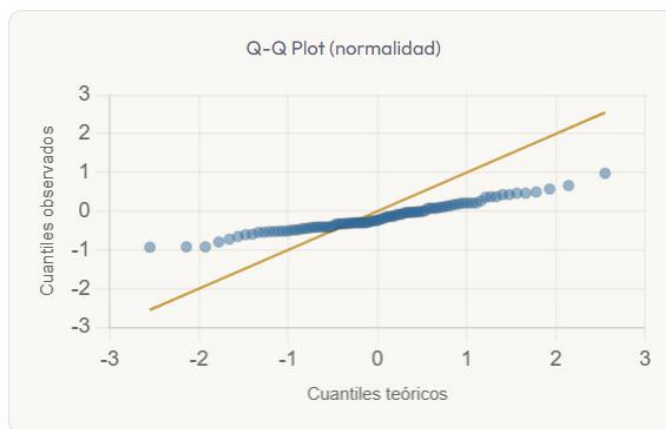
Figura 13. Histograma de residuos (hombres)



Fuente. Elaboración propia mediante procesamiento de datos en Python y visualización en entorno web.

Por su parte, en la figura 13, el histograma de residuos muestra una distribución relativamente centrada, aunque con cierta dispersión, lo que sugiere una aproximación razonable a la normalidad.

Figura 14. Gráfico Q-Q de residuos (hombres)



Fuente. Elaboración propia mediante procesamiento de datos en Python y visualización en entorno web.

Asimismo, en la figura 14, el gráfico Q-Q evidencia que la mayoría de los puntos se alinean cercanamente a la línea teórica de referencia, aunque se observan ligeras desviaciones en los

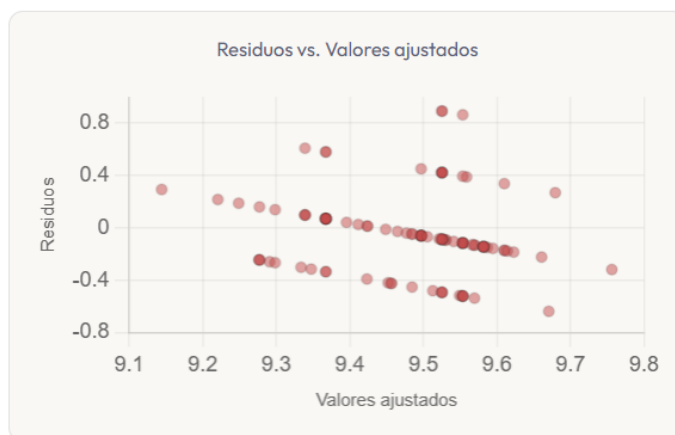
extremos, lo cual puede indicar la presencia de colas ligeramente más pesadas.

Adicionalmente, las estadísticas de los residuos muestran una media de -0.1587 y una desviación estándar de 0.3672, lo cual indica que los errores se encuentran relativamente centrados en torno a cero, aunque con una ligera desviación.

En conjunto, estos resultados sugieren que, si bien el modelo presenta indicios de heterocedasticidad, el supuesto de normalidad se cumple de manera aproximada, lo cual es consistente con los resultados obtenidos en las pruebas estadísticas.

En el caso de las mujeres, se realiza un análisis gráfico de los residuos con el fin de evaluar la adecuación del modelo estimado.

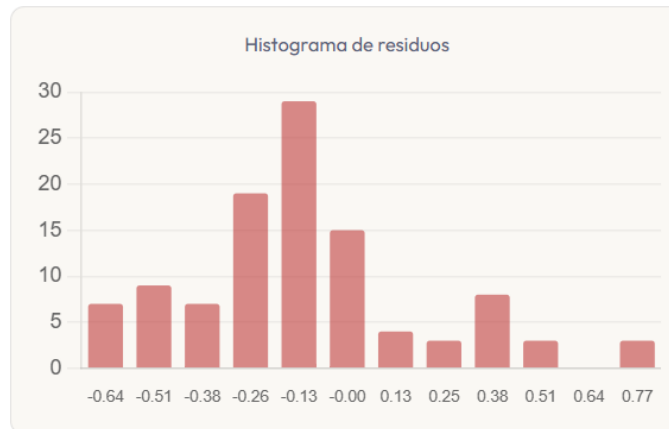
Figura 15. Gráfico de residuos vs. valores ajustados (mujeres)



Fuente. Elaboración propia mediante procesamiento de datos en Python y visualización en entorno web.

En la figura 15, el gráfico de residuos contra valores ajustados, se observa una dispersión relativamente uniforme a lo largo de los valores ajustados, lo cual sugiere la ausencia de patrones sistemáticos claros y es consistente con el cumplimiento del supuesto de homocedasticidad, en línea con los resultados obtenidos en la prueba de Breusch-Pagan.

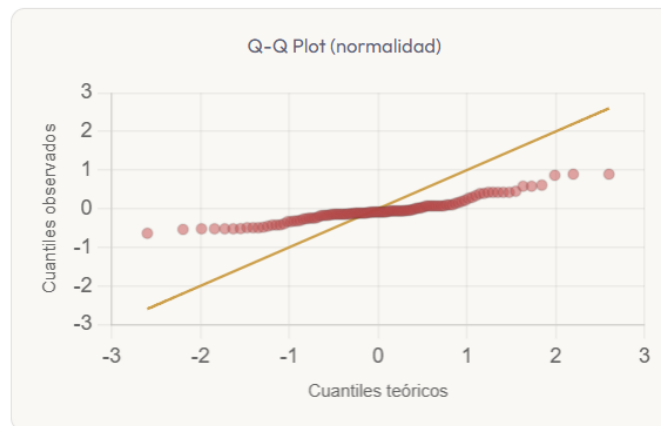
Figura 16. Histograma de residuos (mujeres)



Fuente. Elaboración propia mediante procesamiento de datos en Python y visualización en entorno web.

Por su parte, en la figura 16, el histograma de residuos muestra una distribución centrada, aunque con cierta asimetría, lo que indica que los errores no se distribuyen de manera perfectamente simétrica.

Figura 17. Gráfico Q-Q de residuos (mujeres)



Fuente. Elaboración propia mediante procesamiento de datos en Python y visualización en entorno web.

Asimismo, en la figura 17, el gráfico Q-Q evidencia desviaciones respecto a la línea teórica de referencia, particularmente en los extremos, lo cual sugiere la presencia de colas más pesadas y refuerza la evidencia de no normalidad identificada en la prueba de Jarque-Bera.

Adicionalmente, las estadísticas de los residuos muestran una media de -0.0483 y una desviación estándar de 0.3112, lo cual indica que los errores se encuentran cercanos a cero, aunque con cierta dispersión.

En conjunto, estos resultados sugieren que el modelo cumple adecuadamente con el supuesto de homocedasticidad, pero presenta desviaciones respecto al supuesto de normalidad, lo cual es consistente con los resultados obtenidos en las pruebas estadísticas.

En conjunto, el análisis gráfico de los residuos confirma los resultados obtenidos en las pruebas estadísticas, evidenciando un comportamiento diferenciado entre hombres y mujeres. En el caso de los hombres, se observan indicios de heterocedasticidad, mientras que en el caso de las mujeres la dispersión de los residuos es más uniforme. Sin embargo, en ambos grupos se identifican ligeras desviaciones respecto a la normalidad, particularmente en los extremos de la distribución.

Estos hallazgos refuerzan la importancia de considerar de manera conjunta las pruebas estadísticas y el análisis gráfico, ya que permiten obtener una visión más completa del comportamiento del modelo. Asimismo, justifican la necesidad de interpretar los resultados econométricos con cautela, especialmente en lo que respecta a la inferencia estadística.

4.3.5 Integración de resultados

En conjunto, las pruebas de diagnóstico aplicadas permiten evaluar de manera integral la adecuación del modelo econométrico estimado, evidenciando un comportamiento diferenciado entre hombres y mujeres. En particular, los resultados muestran que, en el caso de los hombres, se identifica la presencia de heterocedasticidad, mientras que el supuesto de normalidad de los residuos se cumple de manera aproximada. Por el contrario, en el caso de las mujeres, no se detectan problemas de heterocedasticidad, pero se observan desviaciones respecto a la normalidad de los errores.

En cuanto a la multicolinealidad, los resultados del factor de inflación de la varianza (VIF) indican que no existen problemas severos en ninguno de los dos grupos, aunque se observa una colinealidad moderada entre las variables de escolaridad y experiencia, la cual resulta

consistente con la teoría del capital humano y no compromete la estabilidad del modelo.

De manera complementaria, el análisis gráfico de los residuos confirma los hallazgos obtenidos en las pruebas estadísticas, mostrando que, si bien el modelo presenta ciertas limitaciones en el cumplimiento de algunos supuestos, estas no invalidan la estimación realizada. En particular, los resultados sugieren que los estimadores obtenidos mediante Mínimos Cuadrados Ordinarios permanecen insesgados, aunque la inferencia estadística debe interpretarse con cautela en aquellos casos donde los supuestos no se cumplen plenamente (Gujarati y Porter, 2009).

En este sentido, los resultados obtenidos permiten concluir que el modelo estimado constituye una herramienta adecuada para el análisis de la brecha salarial, ya que cumple razonablemente con los supuestos fundamentales o presenta desviaciones que son consistentes con la naturaleza de los datos de corte transversal.

A partir de esta validación econométrica, en el siguiente apartado se presentan los resultados de la estimación del modelo y de la descomposición Oaxaca-Blinder, con el objetivo de analizar la magnitud de la brecha salarial y los factores que contribuyen a su explicación dentro de Banca y Seguros.

4.4 Resultados

Una vez validada la especificación del modelo mediante las pruebas de diagnóstico, se procede al análisis de los resultados obtenidos a partir de la estimación del modelo Oaxaca-Blinder. El objetivo de este apartado es cuantificar la magnitud de la brecha salarial entre hombres y mujeres dentro de Banca y Seguros, así como identificar los factores que contribuyen a su explicación.

Con el fin de garantizar la robustez del análisis, se estimaron distintas especificaciones del modelo econométrico, considerando tanto versiones en niveles como en logaritmos, así como diferentes combinaciones de variables explicativas. En particular, se evaluaron modelos que incorporan distintas aproximaciones al capital humano, lo que permitió contrastar la estabilidad de los resultados obtenidos. Los resultados mostraron patrones consistentes en

cuanto a la existencia de una brecha salarial entre hombres y mujeres; sin embargo, tras evaluar las distintas especificaciones estimadas, se observó que el modelo con salario en logaritmos y variables de capital humano ampliado presentó un comportamiento econométrico más estable y consistente en las pruebas de diagnóstico. Por esta razón, dicha especificación fue seleccionada como modelo principal para el análisis y la posterior aplicación de la descomposición Oaxaca-Blinder.

En este contexto, los resultados indican que la brecha salarial total entre hombres y mujeres es de 0.1750 en términos logarítmicos, lo que equivale aproximadamente a una diferencia del 19.1 % en los ingresos promedio, al aplicar la transformación exponencial correspondiente $e^{\beta} - 1$. Este resultado sugiere que, en promedio, las mujeres perciben menores ingresos que los hombres dentro del sector analizado.

Para comprender el origen de esta diferencia, se aplica la descomposición Oaxaca-Blinder, la cual permite separar la brecha salarial en componentes asociados a las características observables y a los retornos a dichas características (Oaxaca, 1973; Blinder, 1973).

En primer lugar, el componente asociado a las diferencias en características observables (dotaciones) presenta un valor de -0.0478, lo que representa aproximadamente -27.34 % de la brecha total. Este resultado indica que, en promedio, las mujeres cuentan con características que, de acuerdo con el modelo, deberían traducirse en mayores niveles de ingreso en comparación con los hombres. En otras palabras, el capital humano no solo limita la explicación de la brecha salarial, sino que opera en sentido contrario.

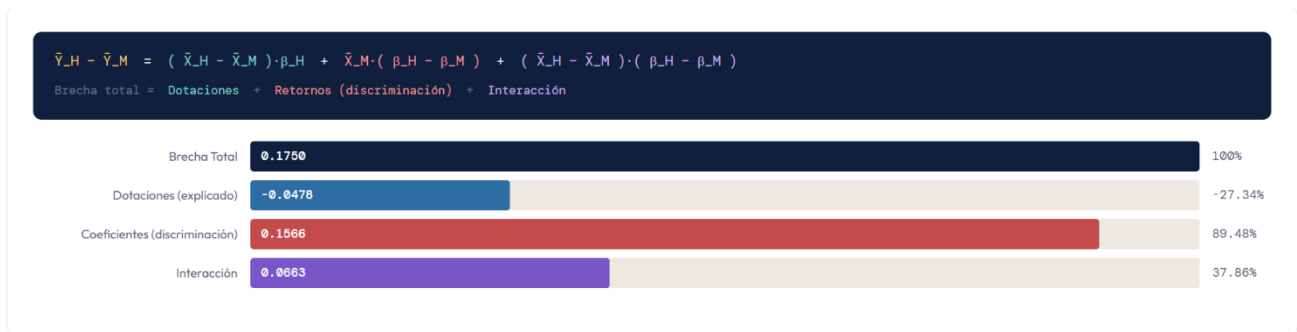
Por el contrario, el componente asociado a los coeficientes, es decir, la parte no explicada de la brecha presenta un valor de 0.1566, lo que representa aproximadamente 89.48 % de la brecha total. Este resultado indica que las diferencias en los retornos a las características laborales favorecen a los hombres, lo que sugiere que el mercado remunera de manera distinta variables como la escolaridad, la experiencia laboral y las horas trabajadas.

Adicionalmente, el componente de interacción presenta un valor de 0.0663, equivalente aproximadamente al 37.86 % de la brecha, lo que refleja un efecto conjunto entre las diferencias en características y en los retornos a las mismas, contribuyendo también a la

desigualdad salarial observada.

En conjunto, estos resultados muestran que la mayor parte de la brecha salarial se concentra en el componente asociado a los coeficientes y su interacción, lo que indica que las diferencias en los retornos a las características observables desempeñan un papel predominante en la desigualdad salarial. Este hallazgo sugiere que la brecha no se origina en diferencias en capital humano, sino en la forma en que el mercado laboral remunera dichas características.

Figura 18. Descomposición de la brecha salarial por componentes (modelo en logaritmos)

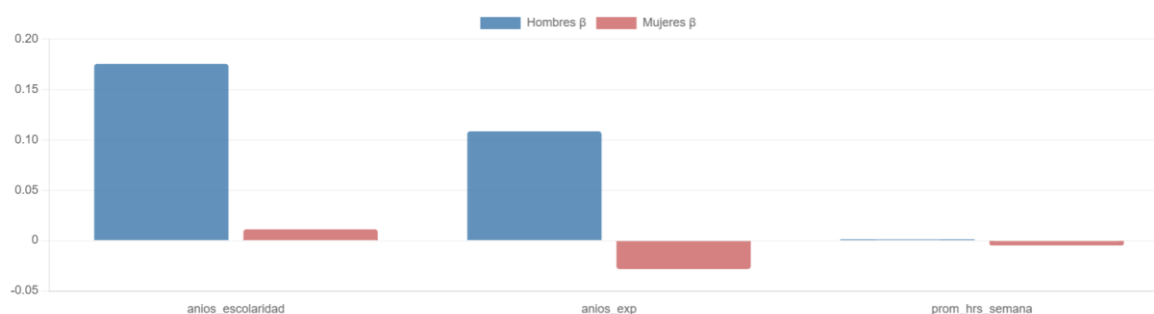


Fuente. Elaboración propia mediante procesamiento de datos en Python y visualización en entorno web.

La Figura 18 presenta la descomposición de la brecha salarial en términos relativos, donde el 100 % corresponde a la brecha total estimada en logaritmos. A partir de esta base, se observa la contribución proporcional de cada uno de los componentes del modelo. En particular, el componente asociado a los coeficientes constituye la mayor proporción de la brecha, seguido del efecto de interacción, mientras que las dotaciones presentan un efecto negativo. Esta representación refuerza la evidencia de que la desigualdad salarial no se explica por diferencias en capital humano, sino por los retornos diferenciados a dichas características.

Al analizar con mayor detalle los coeficientes estimados por grupo, se identifican diferencias significativas en los retornos asociados a las variables de capital humano entre hombres y mujeres.

Figura 19. Coeficientes estimados por sexo (modelo en logaritmos)



Fuente. Elaboración propia mediante procesamiento de datos en Python y visualización en entorno web.

La Figura 19 permite visualizar que la escolaridad y la experiencia laboral presentan efectos positivos y relevantes en el caso de los hombres, mientras que en las mujeres estos retornos son considerablemente menores e incluso negativos en el caso de la experiencia.

Al analizar con mayor detalle los coeficientes estimados por grupo, se identifican diferencias importantes en los retornos asociados a las variables de capital humano. En particular, la escolaridad presenta un efecto positivo considerable en el caso de los hombres, mientras que en las mujeres este efecto es significativamente menor. De manera similar, la experiencia laboral muestra un impacto positivo en los hombres, mientras que en las mujeres incluso presenta un coeficiente negativo, lo que sugiere diferencias estructurales en la forma en que se valoran las trayectorias laborales.

Este comportamiento resulta especialmente relevante, ya que indica que, aun cuando las mujeres poseen características similares o incluso más favorables, estas no se traducen en incrementos proporcionales en sus ingresos. En este sentido, la evidencia sugiere que la brecha salarial no se origina en diferencias en capital humano, sino en la forma en que el mercado laboral remunera dichas características, lo que apunta hacia la existencia de mecanismos estructurales que afectan de manera diferenciada a hombres y mujeres.

Estos resultados son consistentes con la literatura que documenta desigualdades en los retornos al capital humano por género (Blau Y Kahn, 2017) y permiten no solo identificar la existencia de una brecha salarial dentro de Banca y Seguros, sino también comprender su origen. No obstante, si bien estos hallazgos aportan evidencia relevante sobre la estructura de la desigualdad salarial, resulta necesario analizar el alcance del modelo utilizado, así como

sus principales fortalezas y limitaciones. En este sentido, el siguiente apartado presenta una discusión crítica de los resultados, considerando tanto las ventajas del enfoque Oaxaca-Blinder como las restricciones inherentes a su aplicación en este estudio.

4.5 Discusión: ventajas y limitaciones

Los resultados obtenidos a partir de la estimación del modelo Oaxaca-Blinder permiten no solo identificar la existencia de una brecha salarial entre hombres y mujeres dentro de Banca y Seguros, sino también comprender los mecanismos que contribuyen a su formación. En particular, la evidencia empírica sugiere que dicha brecha no se origina en diferencias de capital humano, sino en los retornos diferenciados a estas características. Para interpretar adecuadamente estos hallazgos, resulta pertinente analizar las fortalezas y limitaciones del enfoque adoptado, considerando tanto la calidad de la información utilizada como las características propias de la metodología empleada.

Una de las principales fortalezas de la presente investigación radica en la utilización de la Encuesta Nacional de Ocupación y Empleo (ENOE), elaborada por el Instituto Nacional de Estadística y Geografía (INEGI). Al tratarse de una encuesta oficial con cobertura nacional, diseño probabilístico y representatividad estadística, constituye una de las fuentes más importantes para el estudio del mercado laboral mexicano. Asimismo, la utilización de la información correspondiente al primer trimestre de 2025 permitió disponer del cuestionario ampliado de la ENOE, el cual incorpora información sociodemográfica y laboral más detallada, facilitando la construcción de las variables requeridas para el análisis econométrico.

Otra ventaja importante es la incorporación de factores de expansión, los cuales permiten extrapolar la información observada en la muestra hacia la población objetivo. Esto fortalece la validez de las estadísticas descriptivas y contribuye a que los resultados obtenidos reflejen de manera más adecuada las características de la población analizada.

Desde el punto de vista metodológico, la utilización de la descomposición Oaxaca-Blinder representa una fortaleza relevante, ya que constituye una de las herramientas econométricas más utilizadas para el estudio de brechas salariales. Su principal aportación consiste en

permitir la separación de las diferencias salariales observadas en componentes explicados por características observables y componentes asociados a diferencias en los retornos de dichas características, proporcionando una interpretación más detallada de los mecanismos que generan las desigualdades salariales (Oaxaca, 1973; Blinder, 1973).

Asimismo, la selección de variables se fundamenta en la teoría del capital humano y en las ecuaciones salariales desarrolladas por Mincer (1974), lo que proporciona coherencia entre el marco teórico y el análisis empírico realizado. La inclusión de variables como escolaridad, experiencia laboral potencial y horas trabajadas permite evaluar de manera consistente la contribución de los principales factores tradicionalmente asociados con la determinación de los salarios.

De igual forma, la estimación de diversas especificaciones econométricas, tanto en niveles como en logaritmos del salario, permitió contrastar la estabilidad de los resultados obtenidos. Las pruebas realizadas mostraron resultados cualitativamente consistentes respecto a la existencia de una brecha salarial y a la importancia del componente asociado a los coeficientes, lo que aporta robustez a las conclusiones de la investigación.

Otra fortaleza del estudio se relaciona con la delimitación de la muestra analizada. Al concentrarse en individuos de entre 23 y 27 años que laboran en Banca y Seguros, se reduce parte de la heterogeneidad asociada a diferencias generacionales o etapas avanzadas de la trayectoria laboral. Esto permite analizar con mayor precisión las diferencias salariales presentes en una etapa temprana de la vida profesional, donde los individuos comparten características relativamente similares en términos de experiencia acumulada.

Adicionalmente, la utilización de herramientas de programación en Python permitió automatizar el procesamiento de la información, documentar cada etapa del análisis y garantizar la reproducibilidad de los resultados. Esta característica fortalece la transparencia metodológica y facilita la replicación futura del estudio.

No obstante, también es importante reconocer algunas limitaciones. En primer lugar, la ENOE se basa en información autodeclarada por los individuos, por lo que variables como el ingreso laboral pueden estar sujetas a errores de medición derivados de imprecisiones en el reporte de la información (Bound, Brown Y Mathiowetz, 2001).

En segundo lugar, la interpretación del componente no explicado de la descomposición debe

realizarse con cautela. Aunque frecuentemente se asocia con discriminación salarial u otros factores no observables, la metodología Oaxaca-Blinder no permite establecer relaciones causales directas. En consecuencia, dicho componente puede estar capturando simultáneamente factores relacionados con características individuales no observadas, segmentación ocupacional, calidad del empleo o diferencias institucionales no incluidas en el modelo.

Asimismo, la capacidad explicativa del análisis se encuentra condicionada por la información disponible en la ENOE. Variables potencialmente relevantes, como habilidades específicas, desempeño laboral, mecanismos de promoción interna o características detalladas de los puestos de trabajo, no se encuentran disponibles en la fuente de información utilizada, por lo que no pudieron incorporarse en la estimación.

Finalmente, los resultados deben interpretarse dentro del contexto específico de la población estudiada. Dado que la investigación se concentra en personas de entre 23 y 27 años que laboran en Banca y Seguros, las conclusiones obtenidas no necesariamente pueden generalizarse a otros grupos etarios o sectores económicos sin realizar análisis adicionales.

En conjunto, las fortalezas metodológicas, la calidad de la fuente de información utilizada y la consistencia observada entre las distintas especificaciones estimadas proporcionan evidencia sólida sobre la existencia de una brecha salarial dentro de la población analizada. Si bien las limitaciones señaladas deben considerarse al interpretar los resultados, estas no invalidan los hallazgos obtenidos, sino que delimitan su alcance y abren la posibilidad de futuras investigaciones que profundicen en factores no observados que también pueden contribuir a explicar las diferencias salariales entre hombres y mujeres.

CONCLUSIONES

Principales hallazgos

La presente investigación tuvo como propósito analizar la existencia y los determinantes de la brecha salarial entre hombres y mujeres dentro de Banca y Seguros en México, centrándose en una cohorte específica de trabajadores jóvenes con edades comprendidas entre los 23 y 27 años. Este enfoque permitió reducir la heterogeneidad asociada al ciclo de vida laboral y aislar con mayor precisión los factores que influyen en la formación de las diferencias salariales. A partir de la utilización de la Encuesta Nacional de Ocupación y Empleo (ENOE) y la implementación del modelo de descomposición Oaxaca-Blinder, fue posible no solo identificar la presencia de una brecha salarial, sino también comprender los mecanismos que la explican.

En primer lugar, el análisis descriptivo de la submuestra permitió evidenciar la existencia de una brecha salarial a favor de los hombres dentro del sector analizado. Aunque la magnitud de esta brecha es menor en comparación con la observada a nivel nacional, su persistencia en un grupo relativamente homogéneo resulta particularmente relevante. En este sentido, el hecho de que hombres y mujeres compartan características similares en términos de edad, nivel educativo y experiencia laboral potencial implica que las diferencias salariales no pueden ser atribuidas únicamente a variaciones en el capital humano.

De manera específica, los resultados muestran que tanto hombres como mujeres presentan niveles educativos altamente similares, con una concentración predominante en educación media superior y superior, lo cual es consistente con las características de Banca y Seguros. Asimismo, la experiencia laboral potencial también presenta valores comparables entre ambos grupos. En conjunto, estos hallazgos debilitan la hipótesis tradicional de que las diferencias salariales se explican por disparidades en la formación o experiencia de los individuos.

Adicionalmente, el análisis de la distribución ocupacional permitió identificar la existencia de una segmentación por sexo dentro del sector. Se observó que las mujeres tienden a concentrarse en ocupaciones administrativas, mientras que los hombres presentan una mayor

presencia en ocupaciones comerciales, profesionales y técnicas. Aunque se identificó la presencia de mujeres en puestos de mayor jerarquía, esta participación continúa siendo limitada en términos relativos.

En segundo lugar, la estimación del modelo Oaxaca-Blinder permitió descomponer la brecha salarial en sus distintos componentes. Los resultados muestran que el componente de dotaciones es negativo, lo que implica que, en promedio, las mujeres presentan características iguales o incluso más favorables que los hombres. Sin embargo, el componente de coeficientes resulta positivo y dominante, explicando la mayor parte de la brecha salarial observada.

Este hallazgo sugiere que las diferencias salariales no se originan en lo que las personas poseen en términos de capital humano, sino en la forma en que el mercado laboral valora dichas características. En otras palabras, aun cuando las mujeres cuentan con niveles de educación y experiencia comparables, e incluso ligeramente superiores en algunos casos, estas no se traducen en incrementos proporcionales en sus ingresos.

Por otra parte, la utilización de una especificación logarítmica del salario permitió interpretar los resultados en términos porcentuales, estimando que las mujeres perciben aproximadamente un 19.1 % menos ingresos que los hombres dentro del sector analizado.

En conjunto, los resultados permiten concluir que la brecha salarial de género en Banca y Seguros no puede explicarse únicamente por diferencias en capital humano, sino que responde a una combinación de factores estructurales y de valoración diferencial presentes en el mercado laboral.

Implicaciones y recomendaciones

Los hallazgos obtenidos tienen implicaciones relevantes para las organizaciones de Banca y Seguros. Los resultados muestran que la brecha salarial observada dentro del sector no puede explicarse únicamente por diferencias en educación o experiencia laboral, ya que hombres y mujeres presentan características de capital humano altamente similares. Esto sugiere que continúan existiendo factores estructurales relacionados con la forma en que dichas características son valoradas y remuneradas dentro del mercado laboral.

En este sentido, resulta recomendable que las instituciones del sector fortalezcan sus

mecanismos de evaluación del desempeño, promoción interna y asignación de remuneraciones, con el objetivo de garantizar que las decisiones relacionadas con la compensación económica se encuentren sustentadas en criterios objetivos, transparentes y consistentes. La implementación de revisiones salariales periódicas puede contribuir a identificar posibles diferencias de remuneración que no se encuentren justificadas por factores observables como la escolaridad, la experiencia o las responsabilidades del puesto. Asimismo, los resultados obtenidos evidencian la importancia de continuar promoviendo oportunidades de desarrollo profesional equitativas para hombres y mujeres. La presencia de segmentación ocupacional observada en la muestra sugiere que la distribución de los trabajadores entre distintas categorías laborales sigue siendo un elemento relevante para comprender las diferencias salariales existentes. Por ello, resulta conveniente impulsar estrategias que favorezcan una participación más equilibrada de las mujeres en puestos de mayor responsabilidad, liderazgo y toma de decisiones dentro de las organizaciones. Particularmente, considerando que la presente investigación se centra en trabajadores de entre 23 y 27 años, etapa que generalmente coincide con los primeros años de inserción profesional tras concluir los estudios universitarios, resulta relevante que las organizaciones acompañen el desarrollo de talento desde las fases iniciales de la trayectoria laboral. Programas de mentoría, planes de carrera claramente definidos y procesos de promoción transparentes pueden contribuir a reducir posibles desigualdades que, de mantenerse en el tiempo, tienden a ampliarse conforme avanza la carrera profesional. Finalmente, los resultados permiten señalar que la reducción de la brecha salarial requiere ir más allá del fortalecimiento del capital humano. Si bien la educación y la experiencia continúan siendo factores fundamentales para el desarrollo profesional, la evidencia obtenida muestra que hombres y mujeres con características similares no necesariamente reciben las mismas recompensas económicas dentro del mercado laboral. Por ello, avanzar hacia una mayor igualdad salarial implica no solo fomentar la formación y capacitación de los trabajadores, sino también revisar las prácticas organizacionales mediante las cuales dichas capacidades son reconocidas, valoradas y remuneradas.

Consideraciones para futuras investigaciones

Como fue señalado en el apartado de limitaciones del estudio, la presente investigación se desarrolló a partir de información proveniente de la Encuesta Nacional de Ocupación y Empleo (ENOE), lo que implica ciertas restricciones relacionadas con la disponibilidad de variables y la naturaleza de los datos reportados por los propios individuos. Si bien estos elementos no invalidan los hallazgos obtenidos, sí representan oportunidades para profundizar en el análisis de la brecha salarial de género desde enfoques complementarios.

En primer lugar, futuras investigaciones podrían incorporar variables que no se encuentran disponibles en la ENOE, pero que resultan relevantes para comprender con mayor precisión los mecanismos que generan las diferencias salariales. Entre ellas destacan aspectos como el tipo de puesto desempeñado, el nivel jerárquico dentro de la organización, los esquemas de compensación variable, las evaluaciones de desempeño, las promociones laborales y las responsabilidades de cuidado no remuneradas. La incorporación de estas variables permitiría analizar con mayor detalle cómo se construyen las trayectorias profesionales de hombres y mujeres dentro del sector financiero.

Asimismo, sería de interés ampliar el análisis hacia otras cohortes de edad y etapas del ciclo laboral. La presente investigación se concentró en trabajadores de entre 23 y 27 años, bajo el supuesto de que corresponden principalmente a personas recién egresadas o en las primeras etapas de su trayectoria profesional. Esta decisión permitió analizar una población relativamente homogénea y reducir el efecto de variables asociadas a la experiencia acumulada. Sin embargo, resulta pertinente explorar si las diferencias salariales identificadas se mantienen, disminuyen o se amplían conforme avanza la carrera profesional de los individuos y aumentan sus responsabilidades laborales y familiares.

De igual forma, futuras investigaciones podrían profundizar en el análisis de la segmentación ocupacional dentro del área de Banca y Seguros, particularmente en los mecanismos que limitan o facilitan el acceso de las mujeres a puestos de toma de decisiones, liderazgo y alta dirección. Comprender estas dinámicas permitiría complementar los resultados obtenidos mediante la descomposición Oaxaca-Blinder y aportar evidencia sobre los factores organizacionales que influyen en la persistencia de las desigualdades salariales.

Finalmente, se considera relevante continuar desarrollando estudios que permitan monitorear

la evolución de la brecha salarial de género en el tiempo. En un contexto donde la participación femenina en la educación superior y en sectores altamente especializados ha aumentado de manera significativa, resulta fundamental evaluar si estos avances se traducen efectivamente en igualdad de oportunidades y de remuneración dentro del mercado laboral. La generación constante de evidencia empírica contribuirá no solo al desarrollo académico del tema, sino también al diseño de políticas públicas y estrategias organizacionales orientadas a construir entornos laborales más equitativos.

ANEXOS

ANEXO A. Construcción de variables

Con el propósito de desarrollar el análisis econométrico y la descomposición Oaxaca-Blinder, fue necesario construir, transformar y documentar diversas variables a partir de la información contenida en la Encuesta Nacional de Ocupación y Empleo (ENOE). Algunas variables fueron utilizadas directamente como fueron reportadas por la encuesta, mientras que otras fueron aproximadas mediante proxies, particularmente aquellas relacionadas con capital humano, como la escolaridad y la experiencia laboral potencial.

En este anexo se presentan las principales variables utilizadas en la investigación, así como el procedimiento seguido para la construcción de las variables proxy empleadas dentro de los modelos econométricos.

Tabla A1. Descripción, tipo y construcción de las principales variables empleadas en el análisis econométrico

VARIABLE	DESCRIPCIÓN	TIPO	CONSTRUCCIÓN / ORIGEN
salario_promedio	Ingreso laboral promedio mensual del individuo.	Dependiente	Construcción propia con base en la información de ingresos reportada en la ENOE.
anios_escolaridad	Escolaridad potencial aproximada mediante años de estudio.	Independiente / proxy	Equivalencia en años asignada a partir del nivel educativo categórico reportado en la ENOE.
anios_exp	Experiencia laboral potencial del individuo.	Independiente / proxy	Edad menos años de escolaridad menos seis.
prom_hrs_semana	Promedio de horas trabajadas por semana.	Independiente	Variable construida con base en la información laboral reportada en la ENOE.
sexo_	Sexo del individuo.	Binaria	Codificación utilizada en el análisis: 0 = hombre, 1 = mujer.
e_con	Estado conyugal del individuo.	Categórica	Variable reportada directamente por la ENOE.
entidad	Entidad federativa de residencia.	Categórica	Variable geográfica reportada directamente por la ENOE.
seg_soc	Condición de acceso a seguridad social.	Categórica	Variable reportada directamente por la ENOE.
c_ocu11c	Clasificación de la población ocupada por condición de ocupación.	Categórica	Variable reportada directamente por la ENOE y utilizada para clasificar el tipo de ocupación del individuo.
fac_tri	Factor de expansión trimestral.	Ponderador	Variable reportada por la ENOE para obtener estimaciones representativas a nivel trimestral.

Fuente. Elaboración propia con base en datos de la ENOE.

Tabla A2. Equivalencia del nivel educativo a años de escolaridad

Conversión utilizada para construir la variable de escolaridad potencial a partir de la pregunta 13 de la ENOE.

(CS_P13_1) PREGUNTA 13 ¿HASTA QUÉ GRADO APROBÓ ... EN LA ESCUELA?	NIVEL EDUCATIVO	AÑOS DE ESCOLARIDAD*
1	Preescolar	0
2	Primaria	6
3	Secundaria	9
4	Preparatoria	12
5	Normal	16
6	Carrera técnica	12
7	Profesional	16
8	Maestría	18
9	Doctorado	21
99	No sabe	No considerado en los cálculos

Fuente: Elaboración propia con base en la ENOE. La equivalencia en años fue construida considerando la duración promedio asociada a cada nivel educativo dentro del sistema educativo mexicano. Los registros clasificados como “No sabe” no fueron considerados para la construcción de la variable.

Tabla A3. Clasificación de la población ocupada por condición de ocupación

C_OCU11C	CLASIFICACIÓN DE LA POBLACIÓN OCUPADA POR CONDICIÓN DE OCUPACIÓN
1	Profesionales, técnicos y trabajadores del arte
2	Trabajadores de la educación
3	Funcionarios y directivos
4	Oficinistas
5	Trabajadores industriales, artesanos y ayudantes
6	Comerciantes
7	Operadores de transporte
8	Trabajadores en servicios personales
9	Trabajadores en protección y vigilancia
10	Trabajadores agropecuarios
11	No especificado
0	No especificado

Fuente: Elaboración propia con base en la codificación de la variable **c_ocu11c** reportada en la ENOE.

Fórmula A1. *Construcción de la experiencia laboral potencial*

$$\text{Experiencia laboral potencial} = \text{Edad} - \text{Años de escolaridad} - 6$$

Esta aproximación busca representar el tiempo potencial que un individuo podría haber permanecido dentro del mercado laboral después de concluir su formación académica. Los valores no especificados fueron excluidos de los cálculos.

ANEXO B. Proceso de depuración de datos

Con el propósito de garantizar la consistencia y calidad estadística de la información utilizada en el estudio, fue necesario aplicar un proceso de depuración sobre la base de datos original de la Encuesta Nacional de Ocupación y Empleo (ENOE). Dicho proceso permitió excluir registros con información incompleta, ocupaciones no especificadas y observaciones que no cumplieran con los criterios definidos para la investigación.

La filtración de la información se realiza considerando únicamente individuos pertenecientes a la actividad económica de Banca y Seguros, dentro del rango de edad de 23 a 27 años y con información suficiente para la construcción de las variables utilizadas en el análisis econométrico.

Tabla B1. *Proceso de depuración de la muestra*

ETAPA	CRITERIO APLICADO	OBSERVACIONES
Base original	Registros iniciales de la ENOE antes del proceso de depuración.	423,302
Depuración nacional	Exclusión de registros clasificados como ocupación no especificada y observaciones sin información suficiente para estimar las variables relevantes.	190,503
Selección por edad	Selección de individuos con edades comprendidas entre 23 y 27 años.	21,443
Selección sectorial	Selección de personas ocupadas pertenecientes al sector de Banca y Seguros.	1,899
Muestra final descriptiva	Submuestra final de personas ocupadas en el sector Banca y Seguros con edades entre 23 y 27 años.	269
Muestra final del modelo	Observaciones utilizadas en la estimación del modelo econométrico después de excluir valores faltantes en las variables del modelo.	200

Fuente. Elaboración propia con base en datos de la ENOE.

Tabla B2. Criterios de Inclusión y exclusión aplicados

CRITERIO	DESCRIPCIÓN	DECISIÓN METODOLÓGICA
Ocupación no especificada	Registros sin información laboral suficiente o clasificados en categorías no identificables para el análisis.	Excluidos de la base analítica.
Personas ocupadas	Individuos con condición laboral identificada dentro de la ENOE.	Incluidos cuando contaban con información coherente y suficiente para las variables de interés.
Trabajadores no remunerados	Personas clasificadas como ocupadas, aun cuando no reportaran remuneración monetaria.	Conservados cuando su condición laboral estaba claramente identificada y sus variables eran consistentes.
Sector económico	Personas ocupadas dentro del sector Banca y Seguros.	Seleccionadas para construir la submuestra específica del estudio.
Edad	Individuos entre 23 y 27 años.	Seleccionados para reducir heterogeneidad asociada al ciclo de vida laboral.
Valores faltantes	Observaciones con información faltante en variables clave como salario, escolaridad, experiencia u horas trabajadas.	Excluidas del modelo econométrico cuando impedían la estimación consistente.

Fuente: Elaboración propia con base en los criterios metodológicos definidos para la investigación.

El proceso de depuración permitió construir una base analítica consistente con los objetivos de la investigación, al conservar únicamente observaciones con información suficiente para el análisis descriptivo y econométrico. De esta manera, la muestra final utilizada en el estudio se delimitó a personas ocupadas en la actividad económica de Banca y Seguros, con edades entre 23 y 27 años, lo que permitió reducir la heterogeneidad asociada al ciclo de vida laboral y al sector económico. Esta depuración contribuye a que la estimación del modelo Oaxaca-Blinder se realice sobre una muestra coherente con el planteamiento metodológico de la investigación.

ANEXO C. Resultados complementarios y modelos econométricos

Con el propósito de evaluar distintas especificaciones econométricas y analizar la estabilidad de los resultados obtenidos, se estimaron múltiples modelos utilizando diferentes combinaciones de variables explicativas y transformaciones funcionales. Este proceso permitió comparar el comportamiento estadístico de cada especificación y seleccionar el modelo más adecuado para el análisis de la brecha salarial de género.

En este anexo se presentan tanto las especificaciones utilizadas en los modelos estimados como los principales resultados obtenidos mediante la metodología Oaxaca-Blinder.

Tabla C1. Comparación de especificaciones econométricas estimadas mediante Oaxaca-Blinder

MODELO	TIPO	VARIABLE DEPENDIENTE	VARIABLES INDEPENDIENTES	DESCRIPCIÓN
M1 — Base	NIVEL	salario_promedio	anios_escolaridad, prom_hrs_semana, edad	Modelo base con escolaridad, horas trabajadas y edad.
M2 — Capital humano	NIVEL	salario_promedio	anios_escolaridad, anios_exp, prom_hrs_semana	Modelo centrado en variables de capital humano.
M3 — Estado conyugal	NIVEL	salario_promedio	anios_escolaridad, anios_exp, prom_hrs_semana, e_con	Modelo que incorpora estado conyugal.
M4 — Ocupación	NIVEL	salario_promedio	anios_escolaridad, anios_exp, prom_hrs_semana, c_ocu11c	Modelo que incorpora clasificación ocupacional.
M5 — Entidad	NIVEL	salario_promedio	anios_escolaridad, anios_exp, prom_hrs_semana, entidad	Modelo que incorpora entidad federativa.
M6 — Completo	NIVEL	salario_promedio	anios_escolaridad, anios_exp, prom_hrs_semana, e_con, c_ocu11c, entidad	Modelo ampliado con controles sociodemográficos, ocupacionales y geográficos.
M7 — Seguridad social	NIVEL	salario_promedio	anios_escolaridad, anios_exp, prom_hrs_semana, seg_soc	Modelo que incorpora seguridad social.
M1L — Base log	LOG	ln(salario_promedio)	anios_escolaridad, prom_hrs_semana, edad	Versión semilogarítmica del modelo base.
M2L — Capital humano log	LOG	ln(salario_promedio)	anios_escolaridad, anios_exp, prom_hrs_semana	Modelo principal seleccionado para la investigación.
M3L — Estado conyugal log	LOG	ln(salario_promedio)	anios_escolaridad, anios_exp, prom_hrs_semana, e_con	Versión semilogarítmica con estado conyugal.
M4L — Ocupación log	LOG	ln(salario_promedio)	anios_escolaridad, anios_exp, prom_hrs_semana, c_ocu11c	Versión semilogarítmica con clasificación ocupacional.
M5L — Entidad log	LOG	ln(salario_promedio)	anios_escolaridad, anios_exp, prom_hrs_semana, entidad	Versión semilogarítmica con entidad federativa.
M6L — Completo log	LOG	ln(salario_promedio)	anios_escolaridad, anios_exp, prom_hrs_semana, e_con, c_ocu11c, entidad	Versión semilogarítmica ampliada con controles sociodemográficos, ocupacionales y geográficos.
M7L — Seguridad social log	LOG	ln(salario_promedio)	anios_escolaridad, anios_exp, prom_hrs_semana, seg_soc	Versión semilogarítmica con seguridad social.

Fuente. Elaboración propia con datos de la ENOE mediante modelos Oaxaca-Blinder.

Las especificaciones presentadas en la Tabla C1 incorporan progresivamente distintos controles relacionados con capital humano, características sociodemográficas, ocupación, entidad federativa y seguridad social. Asimismo, se estimaron tanto modelos lineales como semilogarítmicos con el propósito de comparar su capacidad explicativa y su interpretación económica.

Tabla C2. Comparación de especificaciones econométricas estimadas mediante Oaxaca-Blinder

MODELO	TIPO	CONTROLES	Ŷ HOMBRES	Ŷ MUJERES	BRECHA TOTAL	% EXPLICADO	% NO EXPLIC.	% INTERAC.	R ² H	R ² M
M1 — Base	NIVEL	anios_escolaridad, prom_hrs_se...	\$18,178	\$14,323	\$3,855	-16.47%	90.26%	26.21%	0.1542	0.0824
M2 — Cap. humano	NIVEL	anios_escolaridad, anios_exp, pro...	\$18,178	\$14,323	\$3,855	-16.47%	90.26%	26.21%	0.1542	0.0824
M3 — + Est. civil	NIVEL	anios_escolaridad, anios_exp, pro...	\$18,178	\$14,323	\$3,855	-18.06%	58.52%	59.53%	0.2056	0.0828
M4 — + Ocupación	NIVEL	anios_escolaridad, anios_exp, pro...	\$18,178	\$14,323	\$3,855	-9.96%	42.11%	67.85%	0.3004	0.0905
M5 — + Entidad	NIVEL	anios_escolaridad, anios_exp, pro...	\$18,178	\$14,323	\$3,855	-16.47%	90.26%	26.21%	0.1542	0.0824
M6 — Completo	NIVEL	anios_escolaridad, anios_exp, pro...	\$18,178	\$14,323	\$3,855	-12.21%	8.76%	103.45%	0.3394	0.0914
M7 — + Seg. social	NIVEL	anios_escolaridad, anios_exp, pro...	\$18,178	\$14,323	\$3,855	-23.45%	88.46%	34.99%	0.1655	0.0930
M1L — Base (log)	LOG	anios_escolaridad, prom_hrs_se...	9.6813	9.5063	0.1750	-27.34%	89.48%	37.86%	0.1714	0.1027
M2L — Cap. humano (log)	LOG	anios_escolaridad, anios_exp, pro...	9.6813	9.5063	0.1750	-27.34%	89.48%	37.86%	0.1714	0.1027
M3L — + Est. civil (log)	LOG	anios_escolaridad, anios_exp, pro...	9.6813	9.5063	0.1750	-27.67%	61.92%	65.75%	0.2243	0.1035
M4L — + Ocupación (log)	LOG	anios_escolaridad, anios_exp, pro...	9.6813	9.5063	0.1750	-11.62%	49.32%	62.30%	0.3316	0.1133
M5L — + Entidad (log)	LOG	anios_escolaridad, anios_exp, pro...	9.6813	9.5063	0.1750	-27.34%	89.48%	37.86%	0.1714	0.1027
M6L — Completo (log) mejor R²	LOG	anios_escolaridad, anios_exp, pro...	9.6813	9.5063	0.1750	-13.20%	22.25%	90.96%	0.3629	0.1153
M7L — + Seg. social (log)	LOG	anios_escolaridad, anios_exp, pro...	9.6813	9.5063	0.1750	-35.42%	87.09%	48.33%	0.1812	0.1136

Fuente. Elaboración propia con base en estimaciones realizadas a partir de datos de la ENOE mediante modelos Oaxaca-Blinder.

Los resultados muestran que los modelos en logaritmos facilitan la interpretación de la brecha salarial, ya que permiten expresar las diferencias de ingreso de una forma más cercana a porcentajes. Aunque algunos modelos con más variables presentan un mayor R², se seleccionó como modelo principal el M2L — Capital humano (log), porque incluye las variables más importantes para esta investigación: escolaridad, experiencia laboral potencial y horas trabajadas. Esta elección permite mantener un modelo claro y alineado con el objetivo del estudio: identificar qué parte de la brecha salarial se relaciona con diferencias en capital humano y qué parte puede asociarse con diferencias en la forma en que el mercado remunera esas características.

Los resultados complementarios presentados en este anexo permiten transparentar el proceso de selección del modelo econométrico utilizado en la investigación y respaldan

metodológicamente la elección de la especificación final empleada para el análisis de la brecha salarial de género.

ANEXO D. Fragmentos de código utilizados en Python

Con el propósito de garantizar la reproducibilidad y transparencia del análisis, en este anexo se presentan fragmentos del código desarrollado en Python para el procesamiento de datos, construcción de variables, análisis descriptivo y estimación de los modelos econométricos utilizados en la investigación.

Los fragmentos incluidos corresponden a las etapas principales del análisis y permiten ilustrar el procedimiento seguido para la construcción de la base analítica y la implementación de la metodología Oaxaca-Blinder. Cabe señalar que el código completo, junto con la base de datos depurada y los archivos complementarios, se encuentra disponible en el repositorio digital indicado en el Anexo F.

D1. Importación de librerías y bases de datos

```
import pandas as pd #para el manejo de la base de datos
import numpy as np #para operaciones numéricas
import statsmodels.api as sm
from statsmodels.stats.oaxaca import OaxacaBlinder #para estimar regresiones y Oaxaca Blinder
```

Python

```
df = pd.read_excel(r"C:\Users\arlyy\Downloads\brecha_salarial.xlsx") #leer la base de datos
df.head()
```

Python

D2. Preparación y limpieza inicial de variables

```
#convertir variables en numéricas
variables_clave = [
    "sexo_", "eda", "anios_escolaridad", "anios_exp",
    "estado_civil", "salario_promedio", "prom_hrs_semana",
    "seg_soc", "c_oculic", "entidad", "fac_men", "fac_tri"
]

for var in variables_clave:
    print(f"{var}: {'OK' if var in df.columns else 'NO ENCONTRADA'}")

columnas_numericas = [
    "sexo_", "eda", "anios_escolaridad", "anios_exp",
    "salario_promedio", "prom_hrs_semana", "seg_soc",
    "c_oculic", "entidad", "fac_men", "fac_tri"
]

for col in columnas_numericas:
    df[col] = pd.to_numeric(df[col], errors="coerce")
```

Python

```
#crear base de trabajo
df_modelos = df.copy()
```

Python

D3. Recodificación de variables categóricas

```
#recodificar estado civil
def recodificar_estado_civil(x):
    if pd.isna(x):
        return np.nan

    x = str(x).strip().lower().replace("\xa0", " ")

    if x in ["soltero(a)":
        return "soltero"
    elif x in ["casado(a)", "unión libre", "union libre"]:
        return "con_pareja"
    elif x in ["separado(a)", "divorciado(a)", "viudo(a)":
        return "alguna_vez_unido"
    else:
        return "otros"

df_modelos["estado_civil_rec"] = df_modelos["estado_civil"].apply(recodificar_estado_civil)

print(df_modelos["estado_civil_rec"].value_counts(dropna=False))
```

Python

```
#recodificar entidades
freq_ent = df_modelos["entidad"].value_counts(dropna=False)
entidades_validas = freq_ent[freq_ent >= 10].index

df_modelos["entidad_rec"] = np.where(
    df_modelos["entidad"].isin(entidades_validas),
    df_modelos["entidad"],
    "otras_entidades"
)

print(df_modelos["entidad_rec"].value_counts(dropna=False).head(30))
```

Python

```
#creamos dummies
dummies_estado = pd.get_dummies(
    df_modelos["estado_civil_rec"],
    prefix="ec",
    drop_first=True,
    dtype=int)

dummies_ocu = pd.get_dummies(
    df_modelos["c_oculic"],
    prefix="ocu",
    drop_first=True,
    dtype=int)

dummies_ent = pd.get_dummies(
    df_modelos["entidad_rec"],
    prefix="ent",
    drop_first=True,
    dtype=int)

dummies_seg = pd.get_dummies(
    df_modelos["seg_soc"],
    prefix="seg",
    drop_first=True,
    dtype=int)
```

Python

D4. Definición del factor de expansión

```
#factor de expansión
PESO = "fac_tri"
```

Python

D5. Funciones principales del modelo Oaxaca-Blinder

```
def oaxaca_blinder_ponderado(data, y_col, x_cols, group_col, weight_col,
                             grupo_alto=0, grupo_bajo=1):

    d = data[[y_col] + x_cols + [group_col, weight_col]].dropna().copy()

    g1 = d[d[group_col] == grupo_alto].copy()
    g0 = d[d[group_col] == grupo_bajo].copy()

    if len(g1) == 0 or len(g0) == 0:
        raise ValueError("Uno de los grupos no tiene observaciones suficientes.")

    X1 = sm.add_constant(g1[x_cols], has_constant="add")
    X0 = sm.add_constant(g0[x_cols], has_constant="add")

    y1 = g1[y_col]
    y0 = g0[y_col]

    w1 = g1[weight_col]
    w0 = g0[weight_col]

    reg1 = sm.WLS(y1, X1, weights=w1).fit()
    reg0 = sm.WLS(y0, X0, weights=w0).fit()

    xbar1 = pd.Series(
        [1.0] + [np.average(g1[col], weights=w1) for col in x_cols],
        index=["const"] + x_cols)
```

Python

```

)

xbar0 = pd.Series(
    [1.0] + [np.average(g0[col], weights=w0) for col in x_cols],
    index=["const"] + x_cols
)

b1 = reg1.params
b0 = reg0.params

ybar1 = np.average(y1, weights=w1)
ybar0 = np.average(y0, weights=w0)
gap = ybar1 - ybar0

endow = ((xbar1 - xbar0) * b0).sum()
coef = (xbar0 * (b1 - b0)).sum()
inter = ((xbar1 - xbar0) * (b1 - b0)).sum()

detalle = pd.DataFrame({
    "xbar_grupo_alto": xbar1,
    "xbar_grupo_bajo": xbar0,
    "beta_grupo_alto": b1,
    "beta_grupo_bajo": b0,
    "contrib_dotaciones": (xbar1 - xbar0) * b0,
    "contrib_coeficientes": xbar0 * (b1 - b0),
    "contrib_interaccion": (xbar1 - xbar0) * (b1 - b0)
})

```

Python

```

return {
    "gap": gap,
    "endowments": endow,
    "coefficients": coef,
    "interaction": inter,
    "detalle": detalle,
    "modelo_grupo_alto": reg1,
    "modelo_grupo_bajo": reg0,
    "media_y_alto": ybar1,
    "media_y_bajo": ybar0
}

```

Python

```

def porcentajes_oaxaca(resultado):
    gap = resultado["gap"]

    return {
        "pct_explicado": (resultado["endowments"] / gap) * 100,
        "pct_no_explicado": (resultado["coefficients"] / gap) * 100,
        "pct_interaccion": (resultado["interaction"] / gap) * 100
    }

```

Python

```

def correr_modelo_oaxaca(data, y_col, x_cols, nombre_modelo,
                        group_col="sexo_", weight_col=PESO,
                        grupo_alto=0, grupo_bajo=1,
                        exigir_y_positivo=False,
                        exigir_x_positivas=None,
                        verbose=True):

    base = preparar_base_modelo(
        data=data,
        y_col=y_col,
        x_cols=x_cols,
        group_col=group_col,
        weight_col=weight_col,
        grupo_alto=grupo_alto,
        grupo_bajo=grupo_bajo,
        exigir_y_positivo=exigir_y_positivo,
        exigir_x_positivas=exigir_x_positivas,
        verbose=verbose
    )

    resultado = oaxaca_blinder_ponderado(
        data=base,
        y_col=y_col,
        x_cols=x_cols,
        group_col=group_col,
        weight_col=weight_col,

```

Python

```

        grupo_alto=grupo_alto,
        grupo_bajo=grupo_bajo
    )

    pct = porcentajes_oaxaca(resultado)

    resumen = pd.DataFrame([
        "modelo": nombre_modelo,
        "variable_dependiente": y_col,
        "variables_explicativas": ", ".join(x_cols),
        "n_obs": len(base),
        "media_grupo_alto": resultado["media_y_alto"],
        "media_grupo_bajo": resultado["media_y_bajo"],
        "gap_total": resultado["gap"],
        "endowments": resultado["endowments"],
        "coefficients": resultado["coefficients"],
        "interaction": resultado["interaction"],
        "pct_explicado": pct["pct_explicado"],
        "pct_no_explicado": pct["pct_no_explicado"],
        "pct_interaccion": pct["pct_interaccion"]
    ])
    detalle_componentes = resultado["detalle"][[
        "contrib_dotaciones",
        "contrib_coeficientes",
        "contrib_interaccion"
    ]].copy()
    return resultado, resumen, base, detalle_componentes

```

Python

D6. Definición de especificaciones econométricas

```

#Definimos los modelos
x_m1 = ["años_escolaridad", "prom_hrs_semana", "eda"]

x_m2 = ["años_escolaridad", "años_exp", "prom_hrs_semana"]

x_m3 = ["años_escolaridad", "años_exp", "prom_hrs_semana"] + list(dummies_estado.columns)

x_m4 = ["años_escolaridad", "años_exp", "prom_hrs_semana"] + list(dummies_ocu.columns)

x_m5 = ["años_escolaridad", "años_exp", "prom_hrs_semana"] + list(dummies_ent.columns)

x_m6 = ["años_escolaridad", "años_exp", "prom_hrs_semana"] + \
    list(dummies_estado.columns) + list(dummies_ocu.columns) + list(dummies_ent.columns)

x_m7 = ["años_escolaridad", "años_exp", "prom_hrs_semana"] + list(dummies_seg.columns)

```

Python

```

#APLICAR LOGARITMO
df_modelos["ln_salario_promedio"] = np.where(
    df_modelos["salario_promedio"] > 0,
    np.log(df_modelos["salario_promedio"]),
    np.nan
)

```

Python

D7. Estimación del modelo principal

```

#MODELO 2 LOG
resultado_m2_log, resumen_m2_log, base_m2_log, detalle_m2_log = correr_modelo_oaxaca(
    data=df_modelos,
    y_col="ln_salario_promedio",
    x_cols=x_m2,
    nombre_modelo="Modelo 2 - Capital humano ampliado log",
    exigir_y_positivo=False,
    exigir_x_positivas=["prom_hrs_semana"],
    verbose=True
)

print("Gap Total:", resultado_m2_log["gap"])
print("Dotaciones:", resultado_m2_log["endowments"])
print("Coeficientes:", resultado_m2_log["coefficients"])
print("Interacción:", resultado_m2_log["interaction"])

```

Python

D8. Pruebas de diagnóstico

```
#PRUEBAS DE DIAGNÓSTICO

from statsmodels.stats.diagnostic import het_breuschpagan
from scipy.stats import shapiro, jarque_bera

# Guardar regresiones por grupo
reg_hombres = resultado_m2_log["modelo_grupo_alto"]
reg_mujeres = resultado_m2_log["modelo_grupo_bajo"]

# =====
# HETEROCEDASTICIDAD
# =====
bp_hombres = het_breuschpagan(reg_hombres.resid, reg_hombres.model.exog)
bp_mujeres = het_breuschpagan(reg_mujeres.resid, reg_mujeres.model.exog)

print("=== Breusch-Pagan ===")
print("\nHombres")
print("LM Statistic:", bp_hombres[0])
print("LM p-value:", bp_hombres[1])
print("F Statistic:", bp_hombres[2])
print("F p-value:", bp_hombres[3])

print("\nMujeres")
print("LM Statistic:", bp_mujeres[0])
print("LM p-value:", bp_mujeres[1])
print("F Statistic:", bp_mujeres[2])
print("F p-value:", bp_mujeres[3])
```

Python

```
print("F p-value:", bp_mujeres[3])

# =====
# NORMALIDAD
# =====
jb_hombres = jarque_bera(reg_hombres.resid)
jb_mujeres = jarque_bera(reg_mujeres.resid)

print("\n=== Jarque-Bera ===")
print("\nHombres")
print("JB Statistic:", jb_hombres.statistic)
print("p-value:", jb_hombres.pvalue)

print("\nMujeres")
print("JB Statistic:", jb_mujeres.statistic)
print("p-value:", jb_mujeres.pvalue)
```

Python

```
from statsmodels.stats.outliers_influence import variance_inflation_factor
import matplotlib.pyplot as plt
import scipy.stats as stats

# =====
# VIF
# =====
def calcular_vif(modelo):
    X = pd.DataFrame(modelo.model.exog, columns=modelo.model.exog_names)

    vif_df = pd.DataFrame()
    vif_df["variable"] = X.columns
    vif_df["VIF"] = [variance_inflation_factor(X.values, i) for i in range(X.shape[1])]

    return vif_df

vif_hombres = calcular_vif(reg_hombres)
vif_mujeres = calcular_vif(reg_mujeres)

print("=== VIF Hombres ===")
print(vif_hombres)

print("\n=== VIF Mujeres ===")
print(vif_mujeres)
```

```

# =====
# GRAFICOS DE RESIDUOS
# =====
def diagnostico_residuos(modelo, nombre_modelo="Modelo"):
    residuos = modelo.resid
    ajustados = modelo.fittedvalues

    # Residuos vs ajustados
    plt.figure(figsize=(7,5))
    plt.scatter(ajustados, residuos, alpha=0.7)
    plt.axhline(y=0, linestyle="--")
    plt.xlabel("Valores ajustados")
    plt.ylabel("Residuos")
    plt.title(f"{nombre_modelo} - Residuos vs Ajustados")
    plt.show()

    # Histograma
    plt.figure(figsize=(7,5))
    plt.hist(residuos, bins=20)
    plt.xlabel("Residuos")
    plt.ylabel("Frecuencia")
    plt.title(f"{nombre_modelo} - Histograma de residuos")
    plt.show()

    # QQ plot
    plt.figure(figsize=(7,5))

```

Python

Los fragmentos presentados documentan las etapas centrales del procesamiento y análisis. Para efectos de reproducibilidad, el repositorio complementario contiene el notebook completo, la base depurada y los archivos auxiliares utilizados durante la investigación.

ANEXO E. Repositorio digital de materiales complementarios

Con el objetivo de facilitar la consulta y reproducibilidad de los materiales utilizados en esta investigación, se pone a disposición un repositorio digital que concentra la base de datos depurada, scripts utilizados en Python, tablas complementarias, resultados econométricos y materiales gráficos desarrollados durante el análisis.

E1. Repositorio digital

Este repositorio permite consultar los insumos utilizados en el desarrollo de la investigación y contribuye a la transparencia del proceso metodológico seguido.

REFERENCIAS

- Albrecht, J., Björklund, A., y Vroman, S. (2003). ¿Existe un techo de cristal en Suecia? *Journal of Labor Economics*, 21(1), 145–177. <https://doi.org/10.1086/344126>
- Anker, R. (1997). Teorías de la segregación ocupacional por sexo: una revisión general. *International Labour Review*, 136(3), 315–339.
- Asociación Mexicana de Instituciones de Seguros (AMIS). (2024). *Panorama del sector asegurador mexicano*. <https://www.amis.com.mx>
- Banco de México. (1995). *Informe anual 1995*. <https://www.banxico.org.mx/publicaciones-y-prensa/informes-anuales/index.html>
- Banco de México. (2020). *Reporte sobre las economías regionales*.
- Banco de México. (2024). *Reporte sobre las economías regionales*. <https://www.banxico.org.mx/publicaciones-y-prensa/reportes-sobre-las-economias-regionales/index.html>
- Becker, G. S. (1964). *Capital humano: un análisis teórico y empírico, con especial referencia a la educación*. Columbia University Press.
- Blau, F. D., y Kahn, L. M. (2017). La brecha salarial de género: magnitud, tendencias y explicaciones. *Journal of Economic Literature*, 55(3), 789–865. <https://doi.org/10.1257/jel.20160995>
- Blinder, A. S. (1973). Discriminación salarial: estimaciones de forma reducida y estructurales. *Journal of Human Resources*, 8(4), 436–455.
- Borjas, G. J. (2016). *Economía laboral* (7.^a ed.). McGraw-Hill Education.
- Bound, J., Brown, C., y Mathiowetz, N. A. (2001). Error de medición en datos de encuestas. En J. J. Heckman y E. Leamer (Eds.), *Manual de econometría* (Vol. 5, pp. 3705–3843). Elsevier. [https://doi.org/10.1016/S1573-4412\(01\)05012-7](https://doi.org/10.1016/S1573-4412(01)05012-7)
- Breusch, T. S., y Pagan, A. R. (1979). Una prueba simple para heterocedasticidad y variación aleatoria de coeficientes. *Econometrica*, 47(5), 1287–1294. <https://doi.org/10.2307/1911963>
- Cadena Vaca, M. Y. (2020). Discriminación salarial por género: análisis de las empresas del sector privado en el eje central de Bolivia. *Investigación y Desarrollo*, 20(2), 25–40.

Cámara de Diputados del H. Congreso de la Unión. (2024). *Iniciativas y reformas en materia de igualdad salarial y perspectiva de género*. <https://www.diputados.gob.mx>

Comisión Nacional Bancaria y de Valores (CNBV). (2022). *Potencial del sector asegurador*. https://www.gob.mx/cms/uploads/attachment/file/713043/POTENCIAL_DEL_SECTOR_ASEGURADOR_29_03_2022.pdf

Comisión Nacional Bancaria y de Valores (CNBV). (2024). *Boletín estadístico de banca múltiple*. Gobierno de México. <https://www.gob.mx/cnbv>

Comisión Nacional de Seguros y Fianzas (CNSF). (2024). *Anuario estadístico del sector asegurador*. Gobierno de México. <https://www.gob.mx/cnsf>

Comisión Nacional de los Salarios Mínimos (CONASAMI). (2024). *Incremento a los salarios mínimos para 2025*. <https://www.gob.mx/conasami/articulos/incremento-a-los-salarios-minimos-para-2025>

Consejo Nacional de Humanidades, Ciencias y Tecnologías (CONAHCYT). (2020). *Indicadores nacionales de educación superior, ciencia y tecnología*. Gobierno de México. <https://conahcyt.mx>

Constitución Política de los Estados Unidos Mexicanos. (1917). *Artículo 123*. <https://www.diputados.gob.mx/LeyesBiblio/pdf/CPEUM.pdf>

Deloitte. (2024). *Perspectivas del sector financiero en México*. <https://www2.deloitte.com/mx/es.html>

Esquivel, G. (2015). *Desigualdad extrema en México: concentración del poder económico y político*. Oxfam México.

Friedrich-Ebert-Stiftung México. (2011). *El salario mínimo en México: de instrumento de justicia social a obstáculo para el desarrollo*.

Gámez Arroyo, J., Rosas Rojas, E., y Lapa Guzmán, J. (2022). Diferencias salariales por capital humano y por condiciones laborales. *Revista de Estudios en Contaduría, Administración e Informática*, 11(32), 39–60.

Gómez Gallardo, L. S. (2021). Crisis salarial en México: análisis y evolución del salario mínimo y el poder adquisitivo. *Estudios Demográficos y Urbanos*, 36(2), 347–382.

Goldin, C. (2014). Una gran convergencia de género: su último capítulo. *American Economic Review*, 104(4), 1091–1119.

- Greene, W. H. (2012). *Análisis econométrico* (7.^a ed.). Pearson.
- Gujarati, D. N., y Porter, D. C. (2009). *Econometría* (5.^a ed.). McGraw-Hill.
- Hayashi, F. (2000). *Econometría*. Princeton University Press.
- Instituto Mexicano para la Competitividad (IMCO). (2024). *Índice Global de Brecha de Género 2024*. <https://imco.org.mx/indice-global-de-brecha-de-genero-2024/>
- Instituto Nacional de Estadística y Geografía (INEGI). (2005). *Encuesta Nacional de Ocupación y Empleo (ENOE): diseño conceptual*. INEGI.
- Instituto Nacional de Estadística y Geografía (INEGI). (2020). *Encuesta Nacional de Ocupación y Empleo: nueva edición*. <https://www.inegi.org.mx/programas/enoe/15ymas/>
- Instituto Nacional de Estadística y Geografía (INEGI). (2022). *Encuesta Nacional de Ocupación y Empleo Nueva Edición (ENOEN)*. <https://www.inegi.org.mx/programas/enoe/15ymas/>
- Instituto Nacional de Estadística y Geografía (INEGI). (2023). *Encuesta Nacional de Ocupación y Empleo (ENOE)*. <https://www.inegi.org.mx/programas/enoe/15ymas/>
- Jann, B. (2008). La descomposición Blinder-Oaxaca para modelos de regresión lineal. *Stata Journal*, 8(4), 453–479.
- Jarque, C. M., y Bera, A. K. (1987). Una prueba de normalidad para observaciones y residuos de regresión. *International Statistical Review*, 55(2), 163–172. <https://doi.org/10.2307/1403192>
- Ley del Banco de México. (1993). *Diario Oficial de la Federación*. Cámara de Diputados del H. Congreso de la Unión. <https://www.diputados.gob.mx/LeyesBiblio/pdf/LBM.pdf>
- Ley Federal del Trabajo. (1970). *Diario Oficial de la Federación*. Cámara de Diputados del H. Congreso de la Unión. <https://www.diputados.gob.mx/LeyesBiblio/pdf/LFT.pdf>
- Lohr, S. L. (2010). *Muestreo: diseño y análisis* (2.^a ed.). Brooks/Cole.
- Mankiw, N. G. (2014). *Principios de economía* (7.^a ed.). Cengage Learning.
- Marinakakis, A. (2006). *Evolución de los salarios en América Latina*. Organización Internacional del Trabajo.
- Marshall, A. (2013). *Principios de economía*. Síntesis.
- Mincer, J. (1974). *Educación, experiencia e ingresos*. Columbia University Press.
- Moreno-Brid, J. C., y Ros, J. (2009). *Desarrollo y crecimiento de la economía mexicana*:

una perspectiva histórica. Oxford University Press.

Neffa, J. C. (Dir.). (2007). *Teorías económicas sobre el mercado de trabajo II: neoclásicos y nuevos keynesianos*. Fondo de Cultura Económica.

Oaxaca, R. (1973). Diferenciales salariales entre hombres y mujeres en mercados laborales urbanos. *International Economic Review*, 14(3), 693–709.

Organización Internacional del Trabajo (OIT). (2018). *Mujeres y hombres en la economía informal: una perspectiva estadística*.

Organización para la Cooperación y el Desarrollo Económicos (OCDE). (2018). *Panorama de la educación 2018*. <https://www.oecd.org/education/education-at-a-glance/>

Organización para la Cooperación y el Desarrollo Económicos (OCDE). (2021). *Igualdad de género y empoderamiento de mujeres y niñas*. <https://www.oecd.org/gender/>

Organización para la Cooperación y el Desarrollo Económicos (OCDE). (2022). *Panorama de la educación 2022: indicadores de la OCDE*. <https://www.oecd.org/education/education-at-a-glance>

Organización para la Cooperación y el Desarrollo Económicos (OCDE). (2025). *Estudios económicos de la OCDE: México*. OCDE Publishing. <https://www.oecd.org>

ONU Mujeres. (2023). *¿Qué es la brecha salarial?* <https://lac.unwomen.org/es/que-hacemos/empoderamiento-economico/epic/que-es-la-brecha-salarial>

ONU Mujeres. (2025). *Igualdad de género y empoderamiento de las mujeres*. <https://mexico.unwomen.org/es>

Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura (UNESCO). (2020). *Informe de seguimiento de la educación en el mundo 2020: inclusión y educación*. UNESCO. <https://unesdoc.unesco.org>

Pindyck, R. S., y Rubinfeld, D. L. (2018). *Microeconomía* (9.ª ed.). Pearson Educación.

Ríos, H. (2008). *Introducción a la econometría aplicada*. Pearson Educación.

Rodríguez Enríquez, C. (2015). *Economía feminista y economía del cuidado*. Nueva Sociedad. <https://nuso.org/>

Samuelson, P. A., y Nordhaus, W. D. (2010). *Economía con aplicaciones a Latinoamérica* (19.ª ed.). McGraw-Hill.

Schultz, T. W. (1961). Inversión en capital humano. *American Economic Review*, 51(1), 1–

17. <https://www.jstor.org/stable/1818907>

Sistema Nacional de Clasificación de Ocupaciones (SINCO). (2019). *Sistema Nacional de Clasificación de Ocupaciones*. INEGI. <https://www.inegi.org.mx/app/scian/>

Stock, J. H., y Watson, M. W. (2011). *Introducción a la econometría* (3.^a ed.). Pearson.

Varian, H. R. (2014). *Microeconomía intermedia: un enfoque moderno* (9.^a ed.). W. W. Norton y Company.

Banco Mundial. (2020). *Actualización 2020 del índice de capital humano*. <https://www.worldbank.org/en/publication/human-capital>

Wooldridge, J. M. (2010). *Introducción a la econometría: un enfoque moderno* (4.^a ed.). Cengage Learning.

Zurita, B. (2008). La actividad aseguradora: importancia, revisión e integración de conceptos fundamentales. *Investigación Económica*, 67(265), 107–142.