

**UNIVERSIDAD CATÓLICA SANTO TORIBIO DE MOGROVEJO**  
**FACULTAD DE CIENCIAS EMPRESARIALES**  
**ESCUELA DE ECONOMÍA**



**Factores determinantes de la brecha salarial entre hombres y mujeres en el  
mercado laboral peruano: 2014 - 2018**

**TESIS PARA OPTAR EL TÍTULO DE  
ECONOMISTA**

**AUTOR**

**Reyner Ever Torres Coronel**

**ASESOR**

**Antonio Gilberto Escajadillo Durand**

<https://orcid.org/0000-0001-8897-0543>

**Chiclayo, 2023**

**Factores determinantes de la brecha salarial entre hombres y mujeres en el mercado laboral peruano: 2014 - 2018**

**PRESENTADA POR**  
**Reyner Ever Torres Coronel**

A la Facultad de Ciencias Empresariales de la  
Universidad Católica Santo Toribio de Mogrovejo  
para optar el Título de

**Economista**

**APROBADA POR**

Joel Vladimir Diaz Plaza

**PRESIDENTE**

Milagros Carmen Gamarra Uceda

**SECRETARIO**

Antonio Gilberto Escajadillo Durand

**VOCAL**

## **Dedicatoria**

A mis padres y hermanos por su apoyo incondicional brindado.

## **Agradecimientos**

A Dios por brindarme sabiduría y permitirme formar parte de esta distinguida casa de estudios. Así mismo, agradezco a la USAT al igual que a mi asesor de tesis y docentes de todas las asignaturas por hacerme crecer tanto en el ámbito académico como personal.

## INFORME FINAL

---

### INFORME DE ORIGINALIDAD

---

**19%**

INDICE DE SIMILITUD

**18%**

FUENTES DE  
INTERNET

**5%**

PUBLICACIONES

**9%**

TRABAJOS DEL  
ESTUDIANTE

---

### FUENTES PRIMARIAS

---

**1**

[tesis.pucp.edu.pe](http://tesis.pucp.edu.pe)

Fuente de Internet

**2%**

**2**

[repositorio.ucv.edu.pe](http://repositorio.ucv.edu.pe)

Fuente de Internet

**1%**

**3**

[tesis.usat.edu.pe](http://tesis.usat.edu.pe)

Fuente de Internet

**1%**

**4**

Submitted to Pontificia Universidad Catolica del  
Peru

Trabajo del estudiante

**1%**

**5**

[www.consorcio.org](http://www.consorcio.org)

Fuente de Internet

**1%**

**6**

[documentop.com](http://documentop.com)

Fuente de Internet

**1%**

**7**

[uvadoc.uva.es](http://uvadoc.uva.es)

Fuente de Internet

**<1%**

**8**

Submitted to Universidad Rafael Landívar

Trabajo del estudiante

**<1%**

---

## Índice

<b>Resumen .....</b>	<b>9</b>
<b>Abstract .....</b>	<b>10</b>
<b>I. Introducción .....</b>	<b>11</b>
<b>II. Revisión de literatura.....</b>	<b>14</b>
2.1. Antecedentes del problema .....	14
2.2. Bases teóricas científicas .....	19
<b>III. Materiales y métodos .....</b>	<b>23</b>
3.1. Tipo y nivel de investigación .....	23
3.2. Diseño de investigación .....	23
3.3. Población, muestra y muestreo .....	31
3.4. Criterios de selección .....	31
3.5. Operacionalización de variables .....	32
3.6. Técnicas e instrumentos de recolección de datos.....	33
3.7. Procedimientos.....	33
3.8. Plan de procesamiento y análisis de datos .....	34
3.9. Matriz de consistencia.....	36
3.10. Consideraciones éticas .....	37
<b>IV. Resultados y discusión .....</b>	<b>38</b>
4.1. Confiabilidad del instrumento de recolección de datos .....	38
4.2. Análisis descriptivo.....	42
4.3. Análisis inferencial .....	47
4.4. Estimaciones .....	50
<b>V. Conclusiones .....</b>	<b>62</b>
<b>VI. Recomendaciones .....</b>	<b>63</b>
<b>VII. Lista De Referencias.....</b>	<b>64</b>

## Lista de tablas

Tabla 1. Notación simplificada de los términos.....	29
Tabla 2. Métodos de descomposición.....	30
Tabla 3. Operacionalización de variables .....	32
Tabla 4. Matriz de consistencia .....	36
Tabla 5. Coeficiente de correlación de Pearson para el logaritmo natural del ingreso por hora, periodo 2017 – 2018 .....	40
Tabla 6. Coeficiente de correlación de Pearson para la antigüedad en años, periodo 2017 – 2018 .....	41
Tabla 7. Coeficiente de correlación de Pearson para las horas trabajadas, periodo 2017 – 2018	41
Tabla 8. Coeficiente de correlación de Pearson para la edad, periodo 2017 - 2018.....	42
Tabla 9. Distribución de frecuencias de la población de estudio, periodo 2014 – 2018 .....	42
Tabla 10. Medidas de tendencia central y de variabilidad para los hombres, periodo 2014 – 2018 .....	44
Tabla 11. Medidas de tendencia central y de variabilidad para las mujeres, periodo 2014 – 2018 .....	44
Tabla 12. Distribución de los ingresos mensuales por nivel educativo .....	45
Tabla 13. Distribución de los ingresos mensuales por rangos de edad.....	45
Tabla 14. Distribución de los ingresos mensuales según lengua materna .....	46
Tabla 15. Coeficiente de correlación de Pearson: logaritmo del ingreso por hora con las horas trabajadas .....	47
Tabla 16. Coeficiente de correlación de Pearson: logaritmo del ingreso por hora con el nivel de educación .....	47
Tabla 17. Coeficiente de correlación de Pearson: logaritmo del ingreso por hora con la edad	48
Tabla 18. Coeficiente de correlación de Pearson: logaritmo del ingreso por hora con el tamaño de la empresa .....	48
Tabla 19. Coeficiente de correlación de Pearson: logaritmo del ingreso por hora con la antigüedad en la empresa .....	48
Tabla 20. Test de varianza .....	49
Tabla 21. Comparación de medias.....	50
Tabla 22. Test de medias .....	50

Tabla 23. Regresión panel para ambos grupos .....	51
Tabla 24. Regresión panel para los hombres .....	52
Tabla 25. Regresión panel para las mujeres.....	53
Tabla 26. Comparación de las estimaciones de los grupos.....	53
Tabla 27. Interpretación del nivel educativo para hombres y mujeres .....	55
Tabla 28. Interpretación del tamaño de empresa .....	56
Tabla 29. Descomposición de la brecha salarial de género en términos logarítmicos, 2014 – 2018 .....	57

## Lista de figuras

Figura 1. Histograma del logaritmo natural del ingreso por hora, periodo 2014 – 2018.....	39
Figura 2. Distribución normal por cuantiles del logaritmo del ingreso por hora, periodo 2014 - 2018.....	39
Figura 3. Distribución normal por cuantiles de la edad, periodo 2014 -2018 .....	40
Figura 4. Ingreso mensual promedio para hombres y mujeres, periodo 2014 - 2018.....	46
Figura 5. Punto máximo de edad: Hombres, Mujeres y promedio .....	54

## Resumen

Esta investigación tuvo como objetivo analizar los factores determinantes de la brecha salarial entre hombres y mujeres en el mercado laboral peruano: 2014 – 2018. Para lograr dicho objetivo, se usó la base de datos del INEI correspondiente a la ENAHO sobre condiciones de vida y pobreza y se utilizó la encuesta panel para los años correspondientes de la investigación, dentro de esta base de datos se escogió a las personas que respondieron durante estos cinco años. La base de datos fue dividida en dos grupos, una correspondiente a los hombres y la otra a las mujeres para luego ser estimadas por separado mediante el método de Oaxaca – Choe (2016) para panel de datos, los resultados muestran que las mujeres ganan 25.76% menos que los hombres.

Por lo tanto, se encontró que los factores determinantes de la brecha salarial entre grupos son el nivel educativo, las horas trabajadas, la experiencia, el tamaño de la empresa, la edad, la edad elevada al cuadrado. Sin embargo, en el modelo se incluyeron otras variables como es el caso del estado civil, el estar embarazada o no (en el caso de las mujeres), las limitaciones para moverse (discapacidad física), el tiempo de residencia en el distrito o si es perteneciente a un pueblo indígena o no, las cuales no fueron significativas.

**Palabras clave:** brecha salarial, mercado laboral, panel de datos

**Clasificación GEL:** J31, J21, O15, J71, H75, J24, Z22

### **Abstract**

The objective of this research was to analyze the determining factors of the salary gap between men and women in the Peruvian labor market: 2014 - 2018. To achieve this objective, the INEI database corresponding to the ENAHO on living conditions and poverty was used, and the panel survey was increased for the corresponding years of the investigation, within this database the people who responded during these five years were chosen. The database was divided into two groups, one corresponding to men and the other to women, to later be estimated separately using the Oaxaca - Choe (2016) method for data panel, the results show that women earn 25.76 % less than men.

Therefore, it was found that the determining factors of the wage gap between groups are educational level, hours worked, experience, company size, age, age squared. However, other variables are included in the model, such as marital status, being pregnant or not (in the case of women), limitations to movement (physical disability), length of residence in the district or whether belongs to an indigenous people or not, which were not significant.

**Keywords:** wage gap, labor market, data panel

**GEL classification:** J31, J21, O15, J71, H75, J24, Z22

## I. Introducción

En las últimas décadas nuestro país ha evidenciado un constante crecimiento económico y consigo se ha generado un notable incremento al mercado laboral por parte de la mujer, pero esto no se ha visto reflejado en la igualdad de ingresos entre sexos y el principal problema radica en que los hombres perciben un mayor nivel de ingreso que las mujeres. Esto es considerado como una gran injusticia social que genera un alto costo a la economía mundial, según el (Banco Mundial, 2018) encontró que el costo para la economía mundial y para América Latina y el Caribe es de 160 y 6.7 billones de dólares respectivamente.

Como lo señala (Díaz, 2015) la diferencia es notable a pesar de la realización del mismo trabajo o por trabajos de igual valor, esta cuestión termina afectando principalmente a las féminas, independientemente en el sector que trabajen, la carrera profesional que desempeñen, el tipo de contrato que tengan, el tipo de trabajo que realicen o el lugar donde viven.

Por lo que la brecha salarial es definida por Organización Internacional de Trabajo (OIT) como la diferencia del promedio de los ingresos entre hombres y mujeres tomando como base el promedio del ingreso de los hombres.

Así mismo, un estudio elaborado para el Perú por el (Ministerio de Trabajo y Promoción del Empleo – MTPE, 2016) nos dice que el World Economic Forum – WEF fue el primero en presentar el índice global de brecha de género en el año 2006 con la intención de medir las diferencias salariales basadas en el género y al mismo tiempo hacer un seguimiento a la misma. Posteriormente el WEF (2018) mediante el *Global Gender Gap Report* (informe mundial sobre brecha de género) encontró que nuestra región se ubica en el tercer lugar cerrando el 70.8% de la brecha.

Así mismo, la Organización Internacional de Trabajo – OIT en el informe mundial sobre salarios para el periodo 2018 – 2019 sostiene que la brecha salarial entre sexos es más alta en aquellos países donde sus ingresos son bajos y es inferiores en los países de ingresos medios o altos. A nivel de países tenemos a Islandia como el país con la brecha salarial de género más baja y al país con la brecha más alta tenemos a Pakistán.

El (MTPE, 2016) encontró que para este año la brecha de género en el mercado laboral peruano era del 29.2%, esto significa que las trabajadoras peruanas ganaban solo el 70.8% del salario de un hombre. Esta brecha nos ubicaba el puesto 80 de un total de 144 países. Recientemente el (WEF, 2018) situó a nuestro país en el puesto 52 de un total de 149 países, pues para este periodo nos encontrábamos con una brecha cerrada del 72%, si bien es cierto esta brecha ha disminuido en 1.2% en estos dos años, somos conscientes de que esta brecha es

significativa en nuestro país y hay componentes observables y no observables que la determinan.

La (Comisión Europea – CE, 2014) sostiene que la igualdad de salarios entre hombres y mujeres producirá beneficios para la economía, a los empresarios y la sociedad en general. Del mismo modo, al determinar los factores que infieren en las diferencias salariales se tomará medidas que puedan ayudar en reducirlas propiciando un escenario económico más favorable con salarios equitativos. Así mismo la OIT (2013) nos dice que la equidad salarial es una condición fundamental para lograr un trabajo decente, digno y productivo; por otro lado, partiendo del hecho que las mujeres son valoradas y respetadas por sus capacidades, esto ayudará en el mejoramiento en las condiciones en el entorno laboral y servirá a las empresas para potenciar su crecimiento y rentabilidad (Barbosa & Castañeda, 2017).

Hay una extensa literatura sobre la medición de la brecha salarial de género, un estudio elaborado por (Valverdi, 2017) en la cual calcula la brecha salarial y lo divide a está en dos componentes, el primero se refiere a las características propias del individuo y los denomina observables, el segundo tiene que ver con la discriminación al cual lo denomina un factor no observable.

En cuanto a la brecha salarial en nuestro país se dice que ésta puede ser causada por distintos factores, estos pueden ser los factores socio – demográficos, factores de capital humano, factores relacionados con el mercado de trabajo, entre otros. Así mismo, el INEI (2018), señala que la participación de la mujer en el mercado laboral para el año 2017 fue del 44.35% de la Población Económicamente Activa (PEA) de las cuales el 64% de las mujeres se encontraba ocupadas. Si bien es cierto en nuestro país se ha evidenciado una mayor participación de la mujer en el mercado laboral peruano. Entonces, si tanto la inserción laboral de la mujer ha aumentado como el nivel educativo y el ambiente económico ha sido favorable es lógico cuestionarse acerca de la existencia de este problema.

Dado esto, el objetivo principal es responder a la siguiente pregunta de investigación: ¿Cuáles son los factores determinantes de la brecha salarial de género en el mercado laboral peruano: 2014 – 2018? Y de manera empírica poder encontrar resultados que indiquen en qué cuantía los componentes observables y los componentes inobservables de los individuos en la determinación de la brecha salarial.

Como respuesta a ello el objetivo principal de la tesis es identificar los factores determinantes de la brecha salarial entre hombres y mujeres en el mercado laboral peruano: 2014 – 2018. De este modo se desprenden dos objetivos secundarios, y se enumeran de la siguiente manera, i) determinar en qué medida las características propias del individuo generan

las diferencias salariales; ii) determinar en qué medida el factor discriminatorio influye en las diferencias salariales.

Así mismo, la hipótesis que se evidencia junto al objetivo general es si los factores determinantes tienen una relación positiva en la reducción de la brecha salarial entre hombres y mujeres en el mercado laboral peruano: 2014 – 2018.

De tal modo, para identificar los factores que determinan este problema y saber en qué medida las características propias del individuo y el factor discriminatorio influyen en la brecha salarial y para lograr dichos objetivos se utilizará la metodología de Oaxaca – Choe (2016) para datos de panel donde se detallan seis métodos de descomposición que explican la brecha salarial en base a ello se busca propiciar un mercado laboral más equitativo entre hombre y mujeres así mismo se busca reducir el costo social que produce este problema socioeconómico, además, se incrementará la productividad en las empresas potenciando un sostenido crecimiento económico en nuestro país.

La presente investigación tiene una estructura que está dada por la siguiente secuencia. En primer lugar, se buscó relacionar la brecha salarial con los factores que la determinan. En segundo lugar, se ha recurrido a una abundante revisión de la literatura con todo lo relacionado a la brecha salarial, en este caso se encontraron tanto estudios teóricos como empíricos. Encontrando dentro los estudios teóricos a la teoría del capital humano, la teoría de la segmentación de trabajo, la teoría hedonista de los salarios y la teoría de gustos por la discriminación, en cuanto a los estudios empíricos se encontró estudios tanto para el ámbito nacional e internacional. En tercer lugar, y después de hallar evidencia teórica y empírica sobre la brecha salarial se procederá con la elección de nuestro modelo econométrico pertinente para medir la brecha salarial. Para la descomposición de ésta se utilizará la metodología de Oaxaca – Choe (2016) para datos de panel en la cual descomponen a la brecha en seis métodos a estimar con la base de datos panel 2014 – 2018 de la ENAHO; así mismo, todos los procedimientos y análisis de datos serán realizarán en el software STATA 16 y posteriormente se realizará la respectiva corrida del modelo econométrico. cuarto lugar, se presenta los resultados (obtenidos de las regresiones econométricas pertinentes) y discusión. En quinto lugar, se harán las conclusiones de la investigación y finalmente se darán las recomendaciones del investigador.

## II. Revisión de literatura

### 2.1. Antecedentes del problema

La brecha salarial no es ajena entre países, pero hay una clara relación de brecha salarial de género entre los países de ingresos altos y los países de menos ingresos o ingresos bajos. Hay una amplia literatura sobre las brechas salariales de género, entre ellos tenemos:

En un estudio reciente sobre la brecha salarial de género presentado por (T. Denning, Jacob, Lefgren, & vom Lehn, 2019) muestran de qué manera los salarios están relacionados con las horas trabajadas. Entonces, los autores indagan sobre el retorno de las horas trabajadas dentro y entre ocupaciones y de qué manera esto puede explicar a la brecha salarial de género.

Los autores encontraron que si bien es cierto la relación de las horas trabajadas y el salario de un individuo tienen una relación débil, pero el promedio de horas trabajadas dentro de una ocupación está mucho más relacionado con los salarios. Además, esta relación se mantiene, controlando por características individuales y tareas de ocupación. Lo que se encuentra es que un aumento del 10% en las horas trabajadas genera un aumento del 10% a 20% en los salarios según la especificación del trabajo.

Además, se sabe que la relación ha cambiado con el tiempo. Al analizar el periodo 1980 – 2016 vemos que la relación entre las horas trabajadas reales y el salario ha incrementado notablemente, pero en el caso del promedio de las horas ocupacionales y los salarios se ha duplicado en este periodo.

Por otro lado, al existir una relación débil de las horas trabajadas y el salario, este explica muy poco la brecha salarial de género. Finalmente, se encuentra que la brecha salarial de género habría sido mucho menor si el precio de las horas trabajadas no hubiese aumentado con el tiempo, esto significa que el aumento en el retorno a la intensidad ocupacional lo hizo cada vez más difícil para las mujeres lograr la paridad salarial con los hombres.

Por otro lado, en un estudio presentado por (Brindusa, J. Ignacio, & Marra de Artíñano, 2018) tuvo como objetivo determinar la brecha salarial ajustada y la no ajustada de España. Al analizar la brecha salarial de género ajustada se toma en cuenta no solo las variables socioeconómicas como la edad, estado civil, nivel educativo y la experiencia, sino que también se toma en cuenta las características del empleo y la empresa, variables como tamaño de la empresa, tipo de contrato, tipo de jornada, ocupación y sector de actividad. En el caso de la brecha salarial de género sin ajustar se estima mediante Mínimos Cuadrados Ordinarios (MCO) y para la estimación de la brecha salarial de género ajustada se realizan regresiones cuantílicas.

Las estimaciones para ambos métodos fueron realizadas empleando la Encuesta de Estructura Salarial (EES) española y europea. Para el primer caso de la brecha salarial de género sin ajustar se encuentra que esta ha venido decreciendo constantemente, por lo que en el 2002 esta brecha era de 25.3% y para el 2014 esta fue de 17.5%. En cambio, la brecha salarial de género ajustada para el año 2014 fue de 12.7% mientras que en el 2002 fue de 18.7%.

Al compararnos con los países latinos encontramos los siguientes estudios:

(Benítez & Espinoza, 2018) al analizar la discriminación salarial por género en el sector formal ecuatoriano utilizan los registros administrativos, esta base de datos permitió de forma más precisa este fenómeno. En cuanto a la metodología se usó el método de descomposición de Oaxaca – Blinder.

Las variables utilizadas dentro del modelo econométrico fueron: como variable dependiente se tomó el logaritmo del salario por hora, dentro de las variables independientes tenemos a la edad y edad al cuadrado, para el caso del estado civil se utilizó variables dicotómicas para cada categoría, y lo mismo ocurre con el título que tiene la persona, en el caso de que la persona pertenezca al sector público tomo el valor uno y en caso contrario este valor fue cero, para el tamaño de la empresa se utilizó una variable dicotómica para cada categoría, lo mismo ocurrió para identificar las provincias y las ramas de actividad económica.

Lo que hace importante a resaltar del trabajo presentado por estos autores es que el análisis fue realizado mediante las ramas de actividad y tamaño de la empresa, debido al tipo de información con la que se contaba. Los resultados muestran que casi en todas las ramas de actividad existe discriminación, sin embargo, con niveles distintos en cada una de ellas.

Con el objetivo de determinar las diferencias salariales y la discriminación entre sexos por áreas de conocimiento profesional en México y sus regiones según su exposición a la apertura comercial, se desarrolló el estudio planteado por (Rodríguez Pérez & Limas Hernández, 2017). Las áreas de conocimiento tenían la siguiente agrupación dada por la clasificación mexicana de programas de estudio por campos de formación académica para el año 2011 estas tenían la siguiente división: educación; artes y humanidades; ciencias sociales, administración y derecho; ciencias naturales, exactas y de la computación; ingeniería, manufactura y construcción; agronomía y veterinaria, así como salud.

Para este análisis se hizo uso de la Encuesta Nacional de Ocupación y Empleo para el año 2015, tomando como base la metodología de Oaxaca – Blinder (1973) corregido el sesgo de selección, Heckman (1979) y afianzándose del Método DiNardo, Fortin, Lemieux (1996) para su análisis.

La metodología de Oaxaca Blinder y DiNardo, Fortin, Lemicux tienen un parecido en particular donde se realiza un contrafactual simple en la cual solo se consideran los valores medios. A diferencia del primero ésta es una función semi paramétrica que le permite explicar de manera más clara la distribución de los ingresos teniendo en cuenta el efecto que produce el factor sexo, este se explica debido a que trabaja con totalidad con la densidad poblacional.

Los resultados muestran que en el área salud existe una mayor desigualdad y discriminación hacia las mujeres por factores no observables, en el caso de la apertura comercial éste se da en las regiones de alta y baja exposición a la misma. Lo contrario ocurre en el sector educación, donde ellas tienen una mejor posición respecto al resto.

Lo impresionante de esto es encontrar evidencia empírica de que se han adherido al mercado laboral de manera significativa, han accedido a estudios superiores de manera creciente, buen rendimiento académico y en algunas ocasiones han sobrepasado el número de egresos masculinos; pese a esto las mujeres siguen percibiendo un salario inferior al de los hombres que no son atribuibles a las características individuales sino a una posible discriminación en el mercado laboral.

Por otro lado, (Cortés Aguilar & Flórez Vera, 2015) analizan los factores que determinan el ingreso por hora y las diferencias que existe entre géneros para la provincia de Santander de Colombia. Este estudio se toman las variables sobre las características individuales y de capital humano, ya sean estas la educación, experiencia, edad, etc. Las características de empleo como el tipo de ocupación, tamaño de empresa, tipo de trabajo (formal o informal); y, por último, se tomó en cuenta las características familiares, en este caso podemos encontrar variables como estado civil, la tenencia de hijos, etc.

En cuanto a la metodología se utilizó la de Blinder – Oaxaca para evidenciar la existencia del componente discriminatorio, para las estimaciones utilizaron la Gran Encuesta Integrada de Hogares (GEIH) para el periodo 2012 – 2014. Siguiendo a ello, los resultados arrojaron que el 25% y 30% de las diferencias de los ingresos entre sexos se adhiere al componente discriminatorio no observado.

Por otro lado, un estudio elaborado para Argentina por (Galassi & Andrada, 2009) toman como base la ecuación de Mincer basada en la teoría del capital humano, para esta estimación se tomó la base de datos de la Encuesta Permanente de Hogares (EPH) de la segunda mitad del año 2006. Además, se consideró a los jefes de hogar con un número de horas mayores a cero y con ingresos positivos. Este modelo paso por una corrección de sesgo de selección muestral de Heckman; los hallazgos señalan estudiar un año más al igual que la tenencia de hijos y además tener casa propia tiene un efecto negativo en la probabilidad de participar en el mercado laboral,

caso contrario ocurre con la experiencia que tienen un efecto positivo pero decreciente en la probabilidad de participar en el mercado laboral.

El estudio de (Fuentes, Palma, & Montero, 2005) donde analiza la discriminación salarial por género para Chile en la cual se analizó la trascendencia de las diferencias salariales para el periodo 1990 – 2003 utilizando la Encuesta de Caracterización Socioeconómica Nacional (CASEN) elaborada por el Ministerio de Planificación (MIDEPLAN). Los autores emplean dos metodologías con la finalidad de contrastar el hecho de la existencia de una brecha salarial, los modelos que utilizaron fue el de Oaxaca-Blinder (1973) y el modelo de Oaxaca-Ranshem (1994).

El primer modelo señala que las diferencias salariales entre sexos lo determinan básicamente dos componentes, siendo el componente de diferencias del capital humano (dotaciones) y el componente discriminatorio. El segundo modelo tiene un aporte adicional a la del primero, en la cual, cuyo aporte consiste en dividir a la discriminación en discriminación pura que se refiere al subpago a un grupo respecto al otro y favoritismo que consiste en un sobrepago de un grupo respecto al otro.

Como resultado se obtuvo que la brecha salarial ha disminuido en el periodo de estudio, en la cual se encontró una brecha de 27.5% para el año 2003. Esto implica que para este año las mujeres chilenas ganaban solo el 72.5% del salario de los hombres.

Al revisar la literatura nacional encontramos un estudio realizado por el (MTPE, 2017) en la cual realizo un estudio para estimar la brecha de género en el sector privado formal de nuestro país para el periodo de análisis 2004 – 2015 donde emplea dos metodologías para la descomposición de la brecha salarial. En primer lugar, se empleó la metodología de descomposición de Oaxaca (1973) – Blinder (1972) corrigiendo el sesgo de selección, esta metodología consiste en dividir la brecha en las características individuales de los individuos y la discriminación. En segundo lugar, utilizó la metodología de Hugo Ñopo (2004) en la cual no requiere asumir una forma funcional para la ecuación de salarios, por la que se considera un método no paramétrico. Esta metodología se denomina “más allá de media” debido a que se puede observar la contribución de la discriminación en la brecha salarial estimada a lo largo de la distribución del salario, pues con ello se puede hacer un mejor análisis que solamente obtener la brecha en su valor promedio.

Se utilizó la ENAHO del INEI, en la cual se tomó en cuenta a la población con edades de 14 años a más y como se mencionó estas personas se encuentren ocupadas en el sector privado formal. En las cuales se tomó en cuenta las características del individuo, capital humano,

características del hogar y la residencia. En base a ello los resultados obtenidos con estas metodologías fueron las siguientes:

La metodología de Oaxaca – Blinder utilizada para la descomposición de la brecha salarial horaria por género, se encontró que esta ha oscilado entre el 20% y 30% entre el periodo de análisis. En el 2004 la descomposición de la brecha corregida se atribuye de la siguiente manera, el 9% es atribuible a las características individuales, mientras que el 54.8% atribuye a la discriminación salarial, llegando a alcanzar una brecha conjunta de 63.8%. En el 2015, el ingreso de los hombres era superior en 57% a las mujeres, el 10.8% debido a las características individuales y el 46.2% atribuible a la discriminación.

Con la metodología de Ñopo se encuentra que el género influye sobre el salario por hora de manera favorable para los hombres a diferencias de las féminas tomando en cuenta un conjunto de características comparables. Así, en el año 2015, la brecha salario por hora entre hombres y mujeres con similares características fue de 14,4%, de los cuales un 11,1% se asocia al componente discriminatorio. Respecto al año 2004, la brecha salarial habría disminuido 10,7 puntos porcentuales, y a su vez la discriminación salarial habría disminuido 6,4 puntos porcentuales.

Por otro lado, el Banco Central de Reserva del Perú – BCRP (2013) donde analiza las habilidades no cognitivas y brecha salarial de género. El informe se enfoca en el estudio de las habilidades latentes cognitivas y la no cognitivas; para estimar esto se hace uso de dos bases de datos, la primera base de datos es la del Estudio de Niños del Milenio (ENM) del “cohorte joven” y “cohorte adulta” para los años 2002, 2006 y 2009 la segunda base de datos es tomados del Banco Mundial para el año 2010. Por otro lado, nos dice que la posible discriminación podría deberse a la autoselección de ocupaciones que premian distintos tipos de habilidad entre hombres y mujeres.

El modelo estimado por los autores se sustenta en cuatro etapas; en primer lugar, se presenta el modelo de educación, en segundo lugar, el modelo de empleo, en tercer lugar, el modelo de ocupación y finalmente el modelo de salarios. En la cual también se utiliza el método de descomposición ya conocido, que es el de Blinder – Oaxaca. En la que se puede apreciar una significativa brecha salarial de 0.276. Finalmente se concluye que las mujeres perciben un salario inferior a la de los hombres a pesar de que las mujeres tienen una mayor habilidad no cognitiva, esto quiere decir que de algún modo las características que poseen las mujeres son apreciadas negativamente en el mercado laboral y caso contrario ocurre con los hombres donde estas habilidades son premiadas por el empleador o demandante de trabajo.

## 2.2. Bases teóricas científicas

### *Teoría del capital humano*

En esta teoría tenemos a los pioneros del caso como son Porath (1967), Becker (1975) y Mincer (1974), en base a estos modelos se continúan realizando investigaciones hasta nuestros días. Esta teoría considera que los componentes observables en la acumulación de capital humano son la educación y la experiencia, añadiendo que estos componentes son los principales causantes de los ingresos individuales, de las diferencias salariales y de la repartición de la masa salarial (Rabanal, 2001).

Por otro lado, los individuos van a tomar la iniciativa de invertir en capital humano si lo que esperan percibir por esta inversión es mayor a una inversión alternativa. Además, los individuos en sus primeros años de vida designan por completo su tiempo a la educación escolar y toman la decisión de invertir en un estudio superior cuando son jóvenes ya que a medida que aumenta los años también aumentará su costo de oportunidad por la cual esté tendrá menos periodos para recuperar dicha inversión.

Sin embargo, sostiene que la educación y la experiencia no son los únicos que determinan la inversión en el capital humano; sino que la inversión en salud, nutrición, migración y el cuidado que perciben durante los primeros años de vida son parte de ello. Es por ello por lo que una inversión en capital humano aumenta la productividad marginal y consigo su ingreso (Rabanal, 2001).

Mincer (1974) propone un modelo para analizar el efecto que tiene la acumulación del capital humano sobre los ingresos y las diferencias entre ellos, el modelo viene representado de la siguiente manera:

$$\ln(Y) = \beta_0 + \beta_1 \text{Edu} + \beta_2 \text{Exp} + \beta_3 \text{Exp}^2 + \mu$$

Esta ecuación toma como variable dependiente al ingreso expresado en términos logarítmicos y como variables independientes a la educación, la experiencia y la experiencia elevada al cuadrado.

### *Teoría de la segmentación del mercado de trabajo (TSMT)*

Esta teoría nos muestra la cercanía a la aplicación hacia los países subdesarrollados, como es el caso del nuestro. Esta teoría nace en la década de los 60 como una respuesta a la teoría neoclásica y cuenta con cuatro enfoques; en primer lugar, tenemos al enfoque institucionalista de Dunlop 1958 y Kerr 1950, en segundo lugar, el enfoque dualista de Doeringer & Piore 1971,

en tercer lugar, el enfoque marxista de Gordon 1972 y Edwards, Reich & Gordon 1982 y finalmente el enfoque latinoamericano de Souza & Tokman 1976.

Un mercado de trabajo segmentado se da en la situación de que un trabajador con el mismo nivel educativo o con el mismo nivel de productividad percibe un pago diferente en diferentes ocupaciones. Esta teoría sostiene que la segmentación no se genera en el mercado de trabajo, pero sí se manifiesta en éste.

La explicación que se podría dar a esta teoría para los países subdesarrollados difiere en dos puntos principales. En primer lugar; se centra en la insuficiencia del sector moderno para poder absorber toda la fuerza laboral y, en segundo lugar, la existencia de los sindicatos, leyes de salarios mínimos, incentivos para disminuir la rotación de los trabajadores, costos de supervisión o salarios de referencia son causas que parecen reforzar a la segmentación, pero no originarla.

Por otro lado, sostiene que existen dos sectores en el mercado laboral, el primero es el sector moderno y el segundo el sector tradicional. En la cual se puede concluir que la pertenencia a un determinado sector debido al capital es uno de los elementos de las diferencias salariales, así como los mencionados en el párrafo anterior.

En cuanto a las cuestiones salariales, Gimble (1991) concluye que la determinación de salarios no se da solamente por factores económicos, sino que también intervienen los sentimientos y reglas institucionales cambiantes.

### ***Teoría hedonista de los salarios***

La teoría hedonista de los salarios supone que tanto los individuos como los puestos de trabajo son heterogéneos. Esta teoría sostiene que los individuos buscan utilidad (placer) y evitan la desutilidad (dolor), lo que le genera utilidad al individuo sería pues un mejor salario y lo contrario a esto sería la desutilidad que vendrá representada por el riesgo (poca seguridad) en el puesto de trabajo en la cual genera condiciones desagradables para la realización de actividades.

En un análisis experimental elaborado por (Jung, Choe, & Oaxaca, 2016) sostienen que las preferencias por el riesgo también determinan las brechas salariales de género observadas. Este experimento se basa en que los sujetos eligen entre el riesgo, es decir, hablamos de la exposición al desempleo de dichos individuos y, un trabajo seguro después de haberle asignado en rondas tempranas a ambos tipos de trabajos, el detalle es que ambos trabajos implican la misma tarea de escritura. El trabajo arriesgado agrega el elemento de una probabilidad conocida que la oportunidad de escribir no estará disponible en un período determinado.

Los sujetos fueron informados de la prima de riesgo exógena que se ofrece para el trabajo riesgoso. Las mujeres eran más propensas que hombres para seleccionar el trabajo seguro, y estas opciones de trabajo representaron entre el 40% y el 77% de la brecha salarial de género en los experimentos. Por otro lado, las mujeres tenían más aversión al riesgo que los hombres también se manifiesta en los parámetros de aversión al riesgo absoluto constante de Pratt Arrow estimados a partir de una adaptación aleatoria del modelo de utilidad del modelo de cartera de varianza media.

Es por ello, que para algunas empresas el hecho de mantener protegidos y/o asegurados a sus trabajadores resulta más costosa que a otras, esto se debe a que las empresas no laboran en un mismo sector o actividad económica. Además, se sabe que hay personas que prefieren un salario inferior a cambio de una mayor seguridad en el puesto de trabajo, esto implica que existe un emparejamiento entre los trabajadores y los puestos de trabajo al existir una combinación óptima de salario y seguridad. A modo de conclusión de esta teoría se obtiene que las mujeres elegirán racionalmente puestos de trabajo más seguros a pesar de que esto conlleve a ingresos inferiores a la de los hombres.

### ***Modelo del gusto por la discriminación***

Becker sostiene que el discriminador está dispuesto a pagar por la preferencia o gusto a discriminar. Esta teoría es general y uno de sus aspectos que abarca es el hecho de que los empresarios hombres discriminan a las trabajadoras a las cuales les dan empleo, el gusto por discriminar de estos empresarios se basa en que ni ellos ni sus trabajadores quieren trabajar con este grupo.

Si suponemos que el mercado laboral tanto los hombres como las mujeres tienen el mismo nivel de productividad y sus salarios son idénticos, para Becker estos serían sustitutos perfectos y un empresario no discriminador los contratará aleatoriamente. Sin embargo, si un empresario tiene prejuicios sobre la mujer esta situación se invierte y se dice que el empresario tiene gusto por la discriminación y su reacción a esto sería que la contratación de personal femenino le generará costos subjetivos o psíquicos. Este costo subjetivo se puede medir mediante un coeficiente de discriminación que viene representada en unidades monetarias, esto implica que un empresario con prejuicios solo contratará a las mujeres siempre y cuando su salario este por debajo al de los hombres; es decir, el salario de una mujer es el salario que percibe un hombre menos el coeficiente de discriminación.

Entonces podemos afirmar que cuanto mayor sea el gusto por la discriminación por parte del empresario hombre, mayor será las diferencias salariales entre hombres y mujeres en el

mercado laboral. De este modelo podemos concluir que los que salen ganando con la discriminación son los hombres ya que perciben salarios superiores al de las mujeres, por lo que estas terminan perdiendo; por otro lado, los empresarios también terminan perdiendo al pagar un salario mayor por el hecho de contratar mayor cantidad de hombres con el mismo nivel de productividad que las mujeres.

### III. Materiales y métodos

#### 3.1. Tipo y nivel de investigación

Tipo de investigación aplicada porque va a buscarse evidencia empírica de un modelo teórico (Hernandez, Fernandez, & Baptista, 2010). Al considerarse este tipo de investigación no solo se busca aplicar un modelo, sino, encontrar respuestas a múltiples aspectos de mejora en la sociedad.

Además, es de nivel correlacional explicativa porque se centra en explicar el porqué de la ocurrencia de un fenómeno y en qué circunstancias se presenta o por qué se relacionan dos o más variables. Además, pretende establecer las causas de los fenómenos o sucesos que se están estudiando, generan un sentido de entendimiento y son sumamente estructurados; los estudios explicativos se caracterizan porque “están dirigidos a responder por las causas de los eventos y fenómenos físicos o sociales”, (Hernandez, Fernandez, & Baptista, 2010, pág. 95).

#### 3.2. Diseño de investigación

Es cuantitativo porque presenta un conjunto de procesos, es secuencial y probatorio. Es decir, que cada etapa precede de la siguiente y no podemos omitir pasos. El método cuantitativo tiene como finalidad medir las características de los fenómenos sociales y esto se deriva de un marco teórico de acuerdo con el problema planteado que tiene una gama de antecedentes donde se encuentra relaciones entre las variables de estudio, (Bernal, 2010).

No experimental, porque no se alterado o manipulado ningunas de las variables, para su análisis se han considerado dichas variables en su ambiente o estado natural.

El diseño de la investigación es de tipo longitudinal, este tipo de estudios recaban los datos en distintos periodos, para hacer inferencias sobre la evolución del problema, sus causas y sus consecuencias o efectos. Esto supone que el investigador tiene interés por analizar los cambios que se dan a lo largo del tiempo de los distintos acontecimientos.

Es retrospectivo, porque es un estudio que se analiza en el presente, pero tomando como referencia los datos del pasado. En nuestro caso se tomarán los datos de la ENAHO.

De acuerdo con la literatura revisada lo que nos permitirá descomponer la brecha salarial en su parte observada y no observada es la metodología de Oaxaca – Choe para datos de panel, este método fue empleado por primera vez en un estudio realizado para Alemania en el año 2016, en nuestro caso este modelo será aplicado al mercado laboral peruano para determinar las diferencias salariales que existen en el mercado laboral peruano por el hecho de ser hombre o mujer.

Pero debemos señalar previamente que al aplicar esta metodología a la economía peruana los resultados que esperamos con ésta son en una escala muy amplia diferentes debido a la gran diferencia entre economías. Desde una comparativa global entre economías podemos evidenciar que Alemania es una economía muy grande con una fuerza laboral notablemente calificada, alto desarrollo en infraestructura, capital social, un alto nivel de innovación y sobre todo un muy bajo nivel de corrupción. Básicamente su economía se centra en tres sectores, el de mayor tamaño es el sector de servicios que aporta un aproximado del 70%, industria 29.1% y tan solo un 0.9% de agricultura, en porcentajes del Producto interno Bruto (PBI). En nuestro caso, tenemos una economía pequeña y muy heterogénea que tiene un impacto positivo en la generación de la brecha salarial de género.

Nuestro país al estar dedicado netamente a la extracción primaria como la agricultura y la minería hace que se vuelva un poco más expuesto a generar estas brechas; es decir, según la teoría hedonista de los salarios los individuos buscan tener una mayor utilidad, que vendría a ser el placer y evitan la desutilidad, vendría a ser el dolor, en términos económicos se dice que las personas obtienen mayor utilidad cuando obtienen un mayor salario y lo que les genera desutilidad pues será la poca seguridad en el puesto de trabajo. Partiendo de esto y aplicándolo al sector minero podemos ver que el riesgo es alto y por ende la participación de la mujer en este sector es bajo lo que genera una mayor dispersión de salarios entre hombres y mujeres.

En el caso de la agricultura para (Alfaro & Guerrero, 2013) sostiene que en este sector la brecha es más amplia en el segmento de bajos ingresos, por otro lado, se encuentra que el hecho de tener pareja (estar comprometido) impide fuertemente a la mujer en la participación laboral.

De acuerdo con lo anterior se podría decir que la brecha salarial que se espera en nuestro país es sin duda mucho mayor a la brecha salarial encontrada para Alemania, esto podría estar explicada por la heterogeneidad de nuestra estructura económica. Y, por otro lado, es necesario mencionar que nuestra economía es ampliamente informal, que es otro de los factores que nuestro modelo no podría capturar y podría estar explicada como una posible discriminación.

El modelo por estimar se presenta a continuación:

Este modelo econométrico consiste en seis métodos de descomposición con datos de panel de (Oaxaca & Choe, 2016), se presenta mediante un modelo de salarios de la siguiente forma:

$$y_{it} = x_{it}\beta + c_i + u_{it}$$

Donde  $y_{it}$  es la media del salario expresado en términos logaritmos,  $x_{it}$  es un vector  $1 \times K$  de observaciones sobre las covarianzas,  $\beta$  conforma el vector de parámetros  $K \times 1$ ,  $c_i$  es la

heterogeneidad no observada y  $u_{it}$  es el término de error aleatorio. En nuestro caso aplicaremos efectos aleatorios correlacionados, por algún motivo que nuestra base de datos no se encuentre balanceada esté error se alude a dos factores, el primero sería que una parte de nuestra muestra no esté en la PEA (Población Económicamente Activa) y lo segundo estaría relacionado con algún error con la encuesta utilizada.

En este caso, siguiendo a Wooldridge (2010), la media condicional de  $y_{it}$  puede expresarse como:

$$E(y_{it} | x_{it}, \bar{z}_i, \bar{s}_{it} = 1) = x_{it}\beta + \bar{z}_i\pi + \theta_1\lambda_{it} + \theta_2d_{2t}\lambda_{it} + \dots + \theta_Td_{Tt}\lambda_{it}$$

Donde  $s_{it} = 1$  ( $y_{it}$  observado),  $\bar{z}_i$  es un vector  $1 \times J$  de promedios de tiempo individuales para todas las variables exógenas en el modelo incluidas las de  $x_{it}$  y las restricciones de exclusión de la ecuación salarial. Es decir, Para cada individuo los elementos de  $\bar{z}_i$  en cada período se calculan como los promedios de la variable exógena en todos los períodos en los que el individuo aparece en la muestra, no solo los períodos en los que el individuo es empleado.  $J$  representa el número de variables exógenas en el modelo, las variables  $d_{jt}$  son indicadores del periodo,  $T$  es el último periodo de tiempo posible en los datos, en nuestro modelo  $T = 5$  ya que tomaremos datos de cinco años, y  $\lambda_{it}$  es el Ratio de la Inversa de Mills (IMR) asociados con la participación en la fuerza laboral para individuos  $i$  en el periodo  $t$ .

La IMR está expresada de la siguiente manera:

$$\lambda_{it} = \frac{\varphi(\bar{z}_i\hat{\gamma}_t)}{\phi(\bar{z}_i\hat{\gamma}_t)},$$

En este caso vamos a emplear un probit en su forma reducida acerca de la ecuación de selección estimada para un año particular  $t$  tomando en cuenta la variable binaria de la participación de la fuerza laboral  $l_t$  para un  $N_t$  unidades está determinado por:

$$probit(l_{it} = 1 | \bar{z}) = \phi(\bar{z}_i\gamma_t),$$

Donde  $\gamma_t$  es un vector de parámetros  $J \times 1$ .

En esta investigación, las variables IMR serán creadas mediante un modelo probit que será estimado para cada año en particular. Con esto, la IMR que será estimada para cada individuo en cada año será calculada de la siguiente manera:

$$\hat{\lambda}_{it} = \frac{\varphi(\bar{z}_i\hat{\gamma}_t)}{\phi(\bar{z}_i\hat{\gamma}_t)}$$

Por lo tanto, la ecuación final que se tomara en cuenta para la estimación viene representada por:

$$y_{it} = x_{it}\beta + \bar{z}_i\pi + \theta_1\hat{\lambda}_{it} + \theta_2d_{2t}\hat{\lambda}_{it} + \dots + \theta_Td_{Tt}\hat{\lambda}_{it} + error$$

Suponga que la ecuación principal seleccionada de la muestra se estima por separado para hombres y mujeres por Mínimos Cuadrados Ordinarios (MCO), para los hombres la letra representativa será la (h) y para las mujeres esta letra será la (m), entonces:

$$y_{hit} = x_{hit}\hat{\beta}_h + \bar{z}_{hi}\hat{\pi}_h + \hat{\theta}_{h1}\hat{\lambda}_{hit} + \hat{\theta}_{h2}d_{2t}\hat{\lambda}_{hit} + \dots + \hat{\theta}_{hT}d_{Tt}\hat{\lambda}_{hit} + error$$

$$y_{hit} = x_{mit}\hat{\beta}_m + \bar{z}_{mi}\hat{\pi}_m + \hat{\theta}_{m1}\hat{\lambda}_{mit} + \hat{\theta}_{m2}d_{2t}\hat{\lambda}_{mit} + \dots + \hat{\theta}_{mT}d_{Tt}\hat{\lambda}_{mit} + error.$$

En la media general de la muestra de salarios (para todos los periodos de tiempo e individuos), los modelos a estimar vienen dados por:

$$\ddot{y}_h = \ddot{x}_h\beta_h + \ddot{z}_h\hat{\pi}_h + \hat{\theta}_{h1}\ddot{\lambda}_h + \hat{\theta}_{h2}\ddot{\lambda}_{h2} + \hat{\theta}_{h3}\ddot{\lambda}_{h3} + \hat{\theta}_{h4}\ddot{\lambda}_{h4} + \hat{\theta}_{h5}\ddot{\lambda}_{h5}$$

$$\ddot{y}_m = \ddot{x}_m\beta_m + \ddot{z}_m\hat{\pi}_m + \hat{\theta}_{m1}\ddot{\lambda}_m + \hat{\theta}_{m2}\ddot{\lambda}_{m2} + \hat{\theta}_{m3}\ddot{\lambda}_{m3} + \hat{\theta}_{m4}\ddot{\lambda}_{m4} + \hat{\theta}_{m5}\ddot{\lambda}_{m5}$$

$$\text{Donde: } \ddot{y} = \frac{\sum_{i=1}^N \sum_{t=1}^{T_{ei}} y_{it}}{\sum_{i=1}^N T_{ei}}, \quad \ddot{x} = \frac{\sum_{i=1}^N \sum_{t=1}^{T_{ei}} x_{it}}{\sum_{t=1}^N T_{ei}}, \quad \ddot{z} = \frac{\sum_{t=1}^N T_{ei} \bar{z}_i}{\sum_{t=1}^N T_{ei}}, \quad \ddot{\lambda} = \frac{\sum_{i=1}^N \sum_{t=1}^{T_{ei}} \hat{\lambda}_{it}}{\sum_{t=1}^N T_{ei}},$$

$$\ddot{\lambda}_2 = \frac{\sum_{i=1}^N \sum_{t=1}^{T_{ei}} d_{2t} \hat{\lambda}_{it}}{\sum_{t=1}^N T_{ei}} = \frac{\sum_{t=1}^N \hat{\lambda}_{t2}}{\sum_{t=1}^N T_{ei}}, \dots, \ddot{\lambda}_T = \frac{\sum_{i=1}^N \sum_{t=1}^{T_{ei}} d_{Tt} \hat{\lambda}_{it}}{\sum_{t=1}^N T_{ei}} = \frac{\sum_{t=1}^N \hat{\lambda}_{tT}}{\sum_{t=1}^N T_{ei}}, \text{ N es el número de veces}$$

que el  $n$ ésimo individuo aparece en la muestra del salario como empleado o dentro de la PEA.

Cuando se toma en cuenta a la estructura de salarios masculina como base (el salario más alto de los dos grupos) la descomposición de la media general viene determinada por:

$$\begin{aligned} \ddot{y}_h - \ddot{y}_m &= (\ddot{x}_h - \ddot{x}_m)\hat{\beta}_h + (\ddot{z}_h - \ddot{z}_m)\hat{\pi}_h + (\ddot{\lambda}_h - \ddot{\lambda}_m)\hat{\theta}_{h1} + (\ddot{\lambda}_{h2} - \ddot{\lambda}_{m2})\hat{\theta}_{h2} + \dots \\ &\quad + (\ddot{\lambda}_{hT} - \ddot{\lambda}_{mT})\hat{\theta}_{h5} + \ddot{x}_m(\hat{\beta}_h - \hat{\beta}_m) + \ddot{z}_m(\hat{\pi}_h - \hat{\pi}_m) + \ddot{\lambda}_m(\hat{\theta}_{h1} - \hat{\theta}_{m1}) \\ &\quad + \ddot{\lambda}_{h2}(\hat{\theta}_{h2} - \hat{\theta}_{m2}) + \dots + \ddot{\lambda}_{hT}(\hat{\theta}_{hT} - \hat{\theta}_{mT}). \end{aligned}$$

Las diferencias que existe entre la media de las IMR se pueden descomponer en diferencias entre género en los parámetros probit y diferencias de género en los regresores probit, esté se expresa a continuación:

$$\begin{aligned} \ddot{y}_h - \ddot{y}_m &= (\ddot{y}_h - \ddot{y}_m^0) + (\ddot{y}_m^0 - \ddot{y}_h) \\ \ddot{\lambda}_{h2} - \ddot{\lambda}_{m2} &= (\ddot{\lambda}_{h2} - \ddot{\lambda}_{m2}^0) + (\ddot{\lambda}_{m2}^0 - \ddot{\lambda}_{h2}) \\ &\quad \cdot \\ &\quad \cdot \\ &\quad \cdot \\ \ddot{\lambda}_{hT} - \ddot{\lambda}_{mT} &= (\ddot{\lambda}_{hT} - \ddot{\lambda}_{mT}^0) + (\ddot{\lambda}_{mT}^0 - \ddot{\lambda}_{hT}) \end{aligned}$$

$$\text{Donde: } \ddot{y}_m^0 = \frac{\sum_{i=1}^N \sum_{t=1}^{T_{mi}} \hat{\lambda}_{mit}^0}{\sum_{i=1}^{N_m} T_{mi}}, \quad \hat{\lambda}_{mit}^0 = \frac{\varphi(\bar{z}_{mi} \hat{\gamma}_{ht})}{\phi(\bar{z}_{mi} \hat{\gamma}_{ht})}, \quad \ddot{\lambda}_{m2}^0 = \frac{\sum_{i=1}^{N_m} \hat{\lambda}_{mi2}^0}{\sum_{i=1}^{N_m} T_{mi}}, \quad \hat{\lambda}_{mi2}^0 = \frac{\varphi(\bar{z}_{mi} \hat{\gamma}_{h2})}{\phi(\bar{z}_{mi} \hat{\gamma}_{h2})}, \quad \dots,$$

$$\hat{\lambda}_{miT}^0 = \frac{\varphi(\bar{z}_{mi} \hat{\gamma}_{hT})}{\phi(\bar{z}_{mi} \hat{\gamma}_{hT})}, \quad \ddot{\lambda}_{mT}^0 = \frac{\sum_{i=1}^{N_m} \hat{\lambda}_{miT}^0}{\sum_{i=1}^{N_m} T_{mi}}.$$

Los términos  $\ddot{\lambda}_m^0, \ddot{\lambda}_{m2}^0, \dots, \ddot{\lambda}_{mT}^0$  representan la evolución de las IMR de las mujeres usando los parámetros de la estimación probit de los hombres. Tomando en cuenta lo anterior, el término  $(\ddot{\lambda}_h - \ddot{\lambda}_m^0)$  mide cuanta es la diferencia de género en  $(\ddot{\lambda}_h - \ddot{\lambda}_m)$  es atribuible a las diferencia de género en las variables que determinan la selección y  $(\ddot{\lambda}_m^0 - \ddot{\lambda}_m)$  mide cuánto de la diferencia de género surge de esta misma en los parámetros probit en la ecuación de selección. Dichas interpretaciones nos llevan a la siguiente descomposición  $(\ddot{\lambda}_{h2} - \ddot{\lambda}_{m2}), \dots, (\ddot{\lambda}_{h5} - \ddot{\lambda}_{m5})$ .

En la que tenemos a la descomposición más detallada de la siguiente manera:

$$\begin{aligned} \ddot{y}_h - \ddot{y}_m &= (\ddot{x}_h - \ddot{x}_m) \hat{\beta}_h + (\ddot{z}_h - \ddot{z}_m) \hat{\pi}_h + (\ddot{\lambda}_h - \ddot{\lambda}_m^0) \hat{\theta}_{h1}^0 + (\ddot{\lambda}_m^0 - \ddot{\lambda}_m) \hat{\theta}_{h1}^0 \\ &+ (\ddot{\lambda}_{h2} - \ddot{\lambda}_{m2}^0) \hat{\theta}_{h2}^0 + (\ddot{\lambda}_{m2}^0 - \ddot{\lambda}_{m2}) \hat{\theta}_{h2}^0 + \dots + (\ddot{\lambda}_{hT} - \ddot{\lambda}_{mT}) \hat{\theta}_{hT}^0 + (\ddot{\lambda}_{mT}^0 - \ddot{\lambda}_{mT}) \hat{\theta}_{hT}^0 \\ &+ \ddot{x}_m (\hat{\beta}_h - \hat{\beta}_m) + \ddot{z}_m (\hat{\pi}_h - \hat{\pi}_m) + \ddot{\lambda}_m (\hat{\theta}_{h1} - \hat{\theta}_{m1}) + \ddot{\lambda}_{m2} (\hat{\theta}_{h2} - \hat{\theta}_{m2}) + \dots + \ddot{\lambda}_{hT} (\hat{\theta}_{hT} - \hat{\theta}_{mT}). \end{aligned}$$

Cabe señalar que además de esté existen varias formas de los términos de descomposición para poder reflejar tanto las diferencias explicadas y no explicadas, que en otras palabras vendría a ser la discriminación en los salarios entre hombres y mujeres.

Siguiendo a (Oaxaca & Choe, 2016), donde nos muestran seis métodos de descomposición para determinar la parte observada y no observada de la brecha salarial. Por medidas de simplificación de los términos se le ha agregado una notación acotada a cada una de las ecuaciones, la lista de los términos simplificados se encuentra en la Tabla 1.

En el primer método de descomposición se agrupan principalmente las diferencias de género asociadas a las características de la parte explicada de la brecha salarial y las diferencias asociadas con la diferencia de género en los parámetros de la categoría no explicada.

Por otro lado, en el segundo método trata de las diferencias de género en los coeficientes de las IMR como se explica o al menos no es discriminatorio.

En cambio, en el método tercero se da un enfoque algo agnóstico que consiste en identificar un efecto de selección separado que no sea incluido en los componentes explicados o no explicados de la descomposición. El término de selectividad (S3) es resultado de las diferencias de género en los coeficientes de la IMR.

Siguiendo la premisa anterior, en el método cuarto podemos agrupar todas las diferencias de género de las IMR y los coeficientes de la IMR como efectos de la selección. Este método confirma los componentes explicados y no explicados, en primer lugar, de las diferencias de género en ambas covariables que varían en el tiempo y las medias promediadas en el tiempo para los términos de las no IMR y, en segundo lugar, para las diferencias de género en el coeficiente de las covariables que varían en el tiempo y las medias promediadas en el tiempo para los términos de la no IMR.

El método quinto consiste en que todas las diferencias de género en las variables de promedio de tiempo  $\ddot{z}$ , sus efectos aleatorios  $\pi$ , los regresores  $\ddot{x}$  y las diferencias de género en los coeficientes de la IMR. Este método de descomposición elimina el efecto de selección como un componente separado en la descomposición y trata las diferencias de género en los parámetros del probit las ecuaciones de selección como inexplicables. Éste impone el supuesto de que las diferencias de género en heterogeneidad no observada esta capturada por  $\ddot{z}_h \hat{\pi}_h - \ddot{z}_m \hat{\pi}_m$  que son muy diferentes a las diferencias de género en las características observadas que viene determinado por  $(\ddot{x}_h - \ddot{x}_m) \hat{\beta}_h$ .

El método quinto consiste en que todas las diferencias de género en las variables de promedio de tiempo  $\ddot{z}$ , sus efectos aleatorios  $\pi$ , los regresores  $\ddot{x}$  y las diferencias de género en los coeficientes de la IMR. Este método de descomposición elimina el efecto de selección como un componente separado en la descomposición y trata las diferencias de género en los parámetros del probit las ecuaciones de selección como inexplicables. Éste impone el supuesto de que las diferencias de género en heterogeneidad no observada esta capturada por  $\ddot{z}_h \hat{\pi}_h - \ddot{z}_m \hat{\pi}_m$  que son muy diferentes a las diferencias de género en las características observadas que viene determinado por  $(\ddot{x}_h - \ddot{x}_m) \hat{\beta}_h$ .

El método seis, trata las diferencias de género en las variables promediadas en el tiempo  $\ddot{z}$  por completo como mecanismo de selección en el supuesto de que la heterogeneidad no observada esté indisolublemente ligada a la selección.

Los seis métodos de descomposición se muestran de manera resumida en la Tabla 2. En base a ello se harán las estimaciones correspondientes a cada método.

**Tabla 1. Notación simplificada de los términos**

Términos	Notación
$(\ddot{x}_h - \ddot{x}_m)\hat{\beta}_h$	(a)
$(\ddot{z}_h - \ddot{z}_m)\hat{\pi}_h$	(b)
$\ddot{x}_m(\hat{\beta}_h - \hat{\beta}_m)$	(c)
$\ddot{z}_m(\hat{\pi}_h - \hat{\pi}_m)$	(d)
$(\ddot{\lambda}_h - \ddot{\lambda}_m^0)\hat{\theta}_{h1}^0$	(e1)
$(\ddot{\lambda}_m^0 - \ddot{\lambda}_m)\hat{\theta}_{h1}^0$	(e2)
$(\ddot{\lambda}_{h2} - \ddot{\lambda}_{m2}^0)\hat{\theta}_{h2}$	(f1)
$(\ddot{\lambda}_{m2}^0 - \ddot{\lambda}_{m2})\hat{\theta}_{h2}^0$	(f2)
$(\ddot{\lambda}_{h3} - \ddot{\lambda}_{m3}^0)\hat{\theta}_{h3}$	(g1)
$(\ddot{\lambda}_{m3}^0 - \ddot{\lambda}_{m3})\hat{\theta}_{h3}^0$	(g2)
$(\ddot{\lambda}_{h4} - \ddot{\lambda}_{m4}^0)\hat{\theta}_{h4}$	(h1)
$(\ddot{\lambda}_{m4}^0 - \ddot{\lambda}_{m4})\hat{\theta}_{h4}^0$	(h2)
$(\ddot{\lambda}_{h5} - \ddot{\lambda}_{m5}^0)\hat{\theta}_{h5}$	(i1)
$(\ddot{\lambda}_{m5}^0 - \ddot{\lambda}_{m5})\hat{\theta}_{h5}^0$	(i2)
$\ddot{\lambda}_m(\hat{\theta}_{h1} - \hat{\theta}_{m1})$	(j1)
$\ddot{\lambda}_{m2}(\hat{\theta}_{h2} - \hat{\theta}_{m2})$	(j2)
$\ddot{\lambda}_{m3}(\hat{\theta}_{h3} - \hat{\theta}_{m3})$	(j3)
$\ddot{\lambda}_{m4}(\hat{\theta}_{h4} - \hat{\theta}_{m4})$	(j4)
$\ddot{\lambda}_{m5}(\hat{\theta}_{h5} - \hat{\theta}_{m5})$	(j5)

**Tabla 2. Métodos de descomposición**

Métodos de descomposición	Explicada ( $E_i$ )	No explicada ( $U_i$ )	Selectividad ( $S_i$ )
Método 01: $\ddot{y}_h - \ddot{y}_m = E1 + U1$	$a + b + e1 + f1 + g1 + h1 + i1$	$c + d + e2 + f2 + g2 + h2 + i2 + j1 + j2 + j3 + j4 + j5$	
Método 02: $\ddot{y}_h - \ddot{y}_m = E2 + U2$	$a + b + e1 + f1 + g1 + h1 + i1 + j1 + j2 + j3 + j4 + j5$	$c + d + e2 + f2 + g2 + h2 + i2$	
Método 03: $\ddot{y}_h - \ddot{y}_m = E2 + U2 + S3$	$a + b + e1 + f1 + g1 + h1 + i1$	$c + d + e2 + f2 + g2 + h2 + i2$	$j1 + j2 + j3 + j4 + j5$
Método 04: $\ddot{y}_h - \ddot{y}_m = E4 + U4 + S4$	$a + b$	$c + d$	$e1 + e2 + f1 + f2 + g1 + g2 + h1 + h2 + i1 + i2 + j1 + j2 + j3 + j4 + j5$
Método 05: $\ddot{y}_h - \ddot{y}_m = E5 + U5$	$a + b + d + e1 + f1 + g1 + h1 + i1 + j1 + j2 + j3 + j4 + j5$	$c + e2 + f2 + g2 + h2 + i2$	
Método 06: $\ddot{y}_h - \ddot{y}_m = E6 + U6 + S6$	$a$	$c$	$b + d + e1 + e2 + f1 + f2 + g1 + g2 + h1 + h2 + i1 + i2 + j1 + j2 + j3 + j4 + j5$

### **3.3. Población, muestra y muestreo**

La población será tomada a nivel nacional, que corresponde tanto en el área urbana como rural que además abarca los 24 departamentos del país y en la provincia constitucional del Callao.

La muestra está constituida por los hogares encuestados por el Instituto Nacional de Estadística e Informática (INEI) mediante la Encuesta Nacional de Hogares (ENAHO) sobre condiciones de vida y pobreza, ésta se distribuye de la siguiente manera: para el año 2014 fueron 10,011 hogares, en el 2015 fueron 10,950 hogares, en el 2016 fueron 12,164 hogares, en el 2017 fueron 12,038 hogares y para el año 2018 fueron 12,234 hogares.

El tipo de muestreo es probabilístico, por conglomerados. Esta muestra fue dividida de manera aleatoria en 12 submuestras y cada una de ellas con un número aproximado de conglomerados por departamento.

### **3.4. Criterios de selección**

Una vez obtenida la información de los hogares que formaron parte de la ENAHO éstos se tienen que realizar la compatibilidad en todo el periodo y para ello se tomó en cuenta a los hogares que hayan respondido en todos los años del periodo de recolección de datos. Por lo tanto, la muestra panel obtenida durante los años 2014 – 2018 ascienden a los 1,689 hogares compatibles.

Al extraer los datos mediante la ENAHO se sabe que éstos fueron obtenidos mediante el factor básico de muestreo que está determinado por el diseño de la muestra. Por otro lado, para los factores de ajuste por la no respuesta, se ha creído conveniente ajustar los factores teniendo en cuenta la magnitud de la no respuesta, esto se refiere a los rechazos o ausencias.

Los factores de expansión básicos que se tomaron en consideración para la ENAHO 2014 – 2018 fueron ajustados teniendo en cuenta las proyecciones de la población de acuerdo con el componente demográfico.

### 3.5. Operacionalización de variables

Tabla 3. Operacionalización de variables

<i>Tipo de variable</i>	<i>Módulo</i>	<i>Indicadores</i>	<i>Unidad de medida</i>	<i>Escala de medición</i>
Variable dependiente		Logaritmo natural del ingreso por hora	Logaritmo natural del ingreso por hora	Razón
Variables Independientes	Características de los miembros del hogar	Sexo	1=Hombre / 2=Mujer	Nominal
		Estado civil	1=Con compromiso, 2= Sin compromiso - CF 3=Soltero (a)	Nominal
		Edad	Edad en años	Razón
	Educación	Nivel de educación	1=Sin nivel, 2=Primaria, 3=Secundaria, 4= Superior no universitaria, 5=Superior universitaria, 6=Post-Grado	Ordinal
		Lengua materna	1=Castellano / 2=Otra lengua	Nominal
	Salud	Atención de parto	1=Si / 2=No	Nominal
		Limitaciones	1=Si / 2=No	Nominal
	Empleo e Ingresos	Vivía en este distrito hace 5 años	1=Si / 2=No	
		Horas trabajadas	Número de horas trabajadas a la semana	Razón
		Antigüedad	Años de antigüedad en su puesto de trabajo	Razón
		Pueblo indígena	1=Si / 2=No	Nominal
		Tamaño de la empresa	1=Hasta 20 personas, 2=De 21 a 50 personas, 3=De 51 a 100 personas, 4=De 101 a 500 personas, 5=Más de 500 personas	Ordinal
Indicador de PEA		1=Ocupado / 2=Desocupado	Nominal	

### 3.6. Técnicas e instrumentos de recolección de datos

La técnica de recolección de datos es la encuesta que capta la información directamente de una muestra seleccionada, en este caso hemos tomado una encuesta ya elaborada por el INEI, la ENAHO da a conocer las estimaciones de las características demográficas de la población Nacional, Nacional Urbana, Nacional Rural, así como también para la Costa, Sierra y Selva, para el año 2014 – 2018. La ENAHO fue realizada de forma anual para todo el periodo tomado, la muestra panel tomada para el año 2014 fue dividida aleatoriamente en doce submuestras, cada una contenía un número aproximado de conglomerados en cada departamento del país.

El instrumento de recolección de datos que fue empleado es el cuestionario que se centra en obtener información de las múltiples variables.

### 3.7. Procedimientos

Se utilizará la ENAHO PANEL sobre condiciones de vida y pobreza durante el periodo 2014 – 2018 y los hogares de interés para nuestro estudio son aquellos que tienen personas que estén en el rango de 14 – 65 años. Además, se utilizará el módulo dos sobre las características de las personas, el módulo tres de educación, el módulo cuatro sobre salud y el módulo cinco sobre empleo e ingresos ya que en estos módulos se encuentran las variables que se necesitaran para realizar las estimaciones correspondientes.

En todos los módulos empleados las personas tienen un identificador (ID) único dentro de la muestra en la cual está determinada como la variable “numper”, esta variable en algunos módulos no se encuentra ordenada por lo que debemos hacerlo para una correcta fusión de datos entre los módulos.

Teniendo prepara la base de datos panel 2014 – 2018 lo que tenemos que hacer para identificar los datos de la muestra y dividir nuestra base de datos por sexo, se tomará en cuenta una variable binaria, esto implica que en el caso de que el individuo sea hombre se empleará el número uno, en caso contrario será cero. Esta distinción es muy importante para poder obtener el nivel de ingresos por separado, tanto el nivel de salarios para los hombres y para las mujeres.

Entonces, para nuestro análisis se tomará en cuenta el salario por hora de cada individuo, esté se obtendrá dividiendo el salario sobre el total de horas y posteriormente a esta variable se le aplicará logaritmos. Las horas trabajadas representan el total de horas trabajadas del individuo tanto en su ocupación principal como en su ocupación secundaria, se van a excluir a aquellas personas que no trabajen, es decir, a aquellas que sus horas trabajadas son cero.

La variable experiencia se tomará de la variable que nos dice el tiempo (en años) en que el individuo trabaja en esa ocupación.

Finalmente, a las variables estado civil y nivel educativo serán tomadas como variables categóricas, en el caso de la variable estado civil adoptará valores entre 1 y 4 y para el caso del nivel educativo está tomará valores del 1 al 6.

Al contar con todos estos datos obtenidos del ENAHO se procederá a usar regresiones con base de datos longitudinales, en nuestro caso el modelo panel data. Para dar respuesta a la hipótesis y a los objetivos, se analizarán los parámetros de cada variable, también se tendrá en cuenta los niveles de significancia mediante el estadístico t, el cual debe ser mayor al 2%. También, se tendrá en cuenta el  $R^2$  el cual debe ser el mayor posible.

### **3.8. Plan de procesamiento y análisis de datos**

Una vez preparada las variables que nos ayudarán posteriormente en los procedimientos estadísticos y econométricos, procederemos a fusionar los módulos mencionados. Estos módulos serán fusionados en el software estadístico STATA 16 haciendo uso del comando “merge”, para ellos es necesario tener ordenado el ID e insertarlo como el identificador común entre las bases de datos. El fin de utilizar el ID es para añadir las características de un individuo en concreto en los distintos módulos y periodos.

Teniendo nuestra base de datos fusionada; es decir, todos los módulos y los periodos en el mismo archivo. Antes de empezar con el manejo de datos se ha optado por depurar las variables que son ajenas a nuestro análisis, para este procedimiento se puede utilizar cualquiera de los dos comando de STATA que son el comando “drop” y el comando “keep”, por conveniencia se ha usado el comando “keep”, y teniendo en cuenta el tipo de muestreo que emplearemos se procederá a seleccionar una parte de la muestra con las características requeridas para el estudio, es decir, para aquellas personas que respondieron durante todo el periodo de estudio, en nuestro caso para los cinco años. La cuestión está en el cómo identificar a estas personas, esto no es problema que deba preocuparnos ya que en la encuesta está incluida una variable que nos va a ayudar en este proceso, a esta variable se le denomina “perpanel1418” en la cual nos dice: si la persona respondió durante estos cinco años le asigna un valor numérico 1 y 0 en caso contrario. Para ello al software le indicaremos que conserve solamente los datos del primer caso, esto se hace mediante el uso del comando “keep” y se le condiciona con el comando “if”.

Al quedarnos solo con las variables de interés se vio pertinente cambiar la estructura de nuestra base de datos, la reestructuración se llevó a cabo haciendo uso del comando “reshape”, éste comando nos permite cambiar de una estructura vertical o larga o en su lenguaje de STATA “long” a una base horizontal o ancha, en el lenguaje del software “wide”.

Por último, para obtener nuestra base de datos final se van a eliminar todos los espacios vacíos de la variable sexo, se ha optado por este procedimiento ya que nuestro objetivo principal es determinar la brecha salarial entre hombres y mujeres. Al hacer todos los procedimientos pertinentes con la base de datos se obtuvo un total de 4,646 personas compatibles en el panel.

### 3.9. Matriz de consistencia

Tabla 4. Matriz de consistencia

Titulo	Problema	Objetivos	Marco teórico	Hipótesis	Variables	Definición	Dimensiones	Indicadores	Metodología
Factores determinantes de la brecha salarial entre hombres y mujeres en el mercado laboral peruano: 2014 – 2018.	<b>General:</b> ¿Cuáles son los factores determinantes de la brecha salarial de género en el mercado laboral peruano: 2014 – 2018?	<b>General:</b> Analizar los factores determinantes de la brecha salarial de género en el mercado laboral peruano: 2014 – 2018.	Modelo a aplicar: $y_{it} = x_{it}\beta + c_i + u_{it}$	<b>General:</b> Los factores determinantes tienen una relación positiva en la reducción de la brecha salarial entre hombres y mujeres en el mercado laboral peruano: 2014 – 2018.	<b>Variable dependiente</b>  Logaritmo del ingreso por hora	Son los ingresos que perciben los individuos por la contraprestación de servicios, aplicada en logaritmos.	Económica	Nivel de ingresos	<b>Tipo:</b> Aplicada y explicativa  <b>Método:</b> Cuantitativo  <b>Diseño de contraste de hipótesis:</b> No experimental, de corte longitudinal y retrospectivo
	<b>Específicos:</b>	<b>Específicos:</b> • Determinar en qué medida las características propias del individuo influye en la determinación de la brecha salarial.		<b>Específicos</b>	Variables independientes • Horas trabajadas  • Nivel educativo  • Experiencia y experiencia al cuadrado	<ul style="list-style-type: none"> <li>Se entiende por horas trabajadas el número total de horas efectivamente trabajadas por el personal ocupado durante el año.</li> <li>Nivel de educación más alto que una persona ha terminado.</li> <li>La acumulación de conocimientos prácticos que una persona ha adquirido en el desempeño de sus funciones</li> </ul>	Tiempo  • Social  • Social	<ul style="list-style-type: none"> <li>Horas</li> <li>Categoría: del 1 al 6.</li> <li>Años</li> </ul>	$y_{it} = x_{it}\beta + c_i + u_{it}$
		• Determinar en qué medida el factor discriminatorio influye en la determinación de la brecha salarial.			• Edad  • Estado civil  • Tamaño de la empresa	<ul style="list-style-type: none"> <li>Tiempo que ha vivido una persona u otro ser vivo contando desde su nacimiento.</li> <li>Condición de una persona según el registro civil en función de si tiene o no pareja y su situación legal respecto a esto</li> <li>expresar el mayor o menor volumen de la empresa.</li> </ul>	• Social  • Social  • Económica	<ul style="list-style-type: none"> <li>Años</li> <li>Categoría: del 1 al 5.</li> <li>Categoría: del 1 al 5.</li> </ul>	<b>Población:</b> Perú <b>Muestra:</b> hogares: 2014 – 2018 <b>instrumentos de recolección de datos:</b> Encuesta sobre las condiciones de vida y pobreza de la ENAHO PANEL

### **3.10. Consideraciones éticas**

No se ha alterado ninguna base de datos ni se ha difundido información privada de los encuestados por el INEI.

## IV. Resultados y discusión

### 4.1. Confiabilidad del instrumento de recolección de datos

Para determinar la confiabilidad del instrumento se realiza el método de estabilidad con la prueba Test – Retest. Entonces, se considera que un instrumento es válido cuando se obtiene resultados similares y se correlacionan entre sí en dos periodos distintos aplicado al mismo objeto de estudio.

Para evaluar el test ya mencionado se hace uso de dos coeficientes de correlación, estos son el coeficiente de correlación de Pearson y el de Spearman. Para saber cuál aplicar se debe probar la normalidad de las variables. Por otro lado, para probar la normalidad se deben cumplir tres criterios: las variables tienen que ser cuantitativas, simétrica y asintótica. En base a esto se debe determinar la normalidad de las variables, para ello se utiliza a Shapiro-Wilk y Kolmogórov-Smirnov. En este caso es más sencillo saber que usar, consiste simplemente en observar al tamaño de la muestra, en caso de que tengamos menor a 50 se utiliza el primero y en caso de que tengamos una muestra mayor o igual a 50 se utilizará el segundo.

Para poder hacer las interpretaciones de manera correcta se plantean dos hipótesis:

$H_0$  : la muestra cuenta con una distribución normal.

$H_1$  : la muestra cuenta con una distribución no normal.

Ahora, para poder aceptar o rechazar las hipótesis planteadas se debe cumplir con los siguientes criterios, teniendo un nivel de significancia del 95%.

$p < 0.05$  : se rechaza la hipótesis nula, aceptando que la muestra cuenta con una distribución no normal.

$p \geq 0.05$  : se acepta la hipótesis nula, aceptando que la muestra cuenta con una distribución normal.

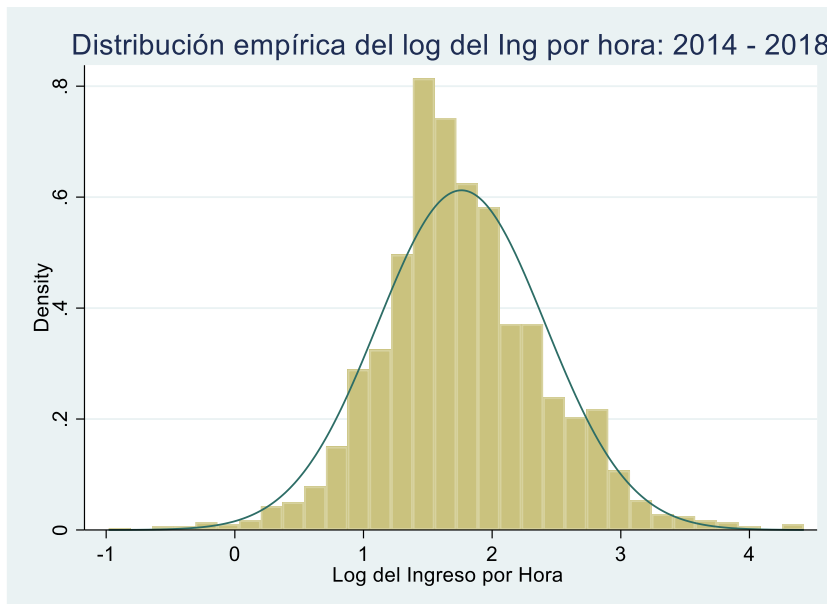
Teniendo prácticamente todo lo necesario para realizar este tipo de pruebas, en este trabajo no se realizarán dichas pruebas, ya que al trabajar con una muestra significativamente grande se asume una distribución normal; sin embargo, se llegará a demostrar la normal de las variables mediante análisis gráfico.

En la Figura 1, podemos observar que nuestra dependiente (el logaritmo natural de los ingresos por hora) se asemeja a una distribución normal.

Otro tipo de gráficos que podemos analizar y que también resulta sencillo de interpretar es el gráfico de cuantiles de distribución normal. Este tipo de gráfica nos indica que aquellos puntos

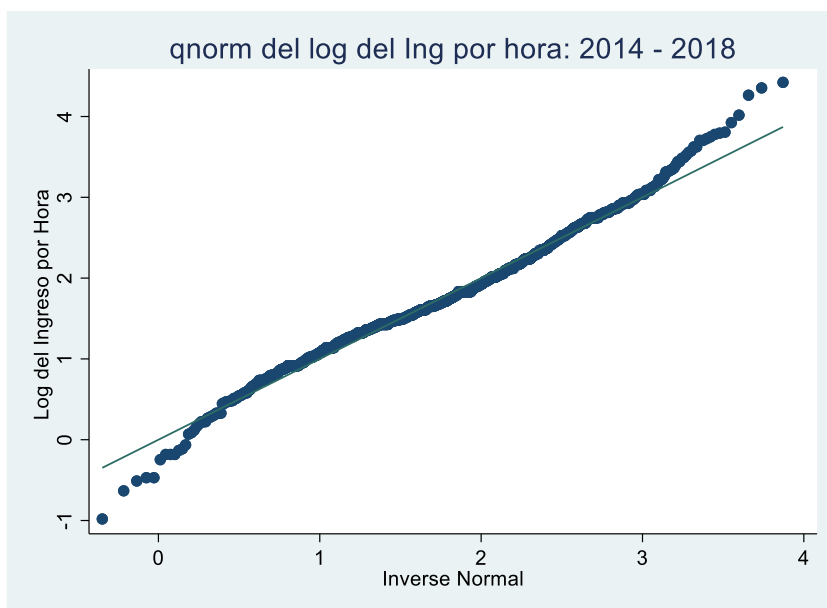
que chocan o están pegadas a la línea con pendiente positiva tienen una distribución normal o tienden a tener una distribución de este tipo.

**Figura 1. Histograma del logaritmo natural del ingreso por hora, periodo 2014 – 2018**



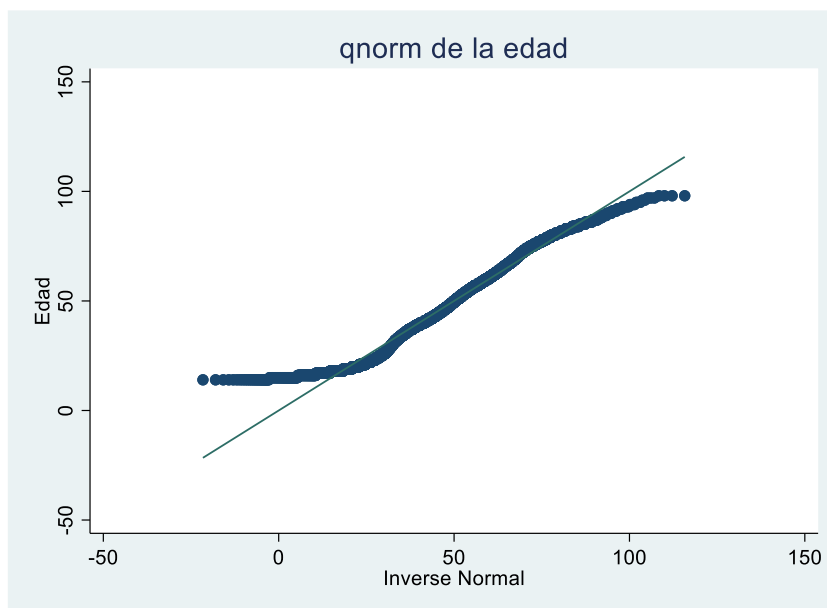
Otra variable cuantitativa que tenemos en nuestro modelo econométrico es la edad, esta distribución lo podemos ver en la Figura 3.

**Figura 2. Distribución normal por cuantiles del logaritmo del ingreso por hora, periodo 2014 - 2018**



Lo que se puede observar en la Figura 3, es que los puntos están muy pegados a la línea, por lo que se dice que tiene una distribución normal.

**Figura 3. Distribución normal por cuantiles de la edad, periodo 2014 -2018**



Una vez demostrada la normalidad de las variables, se procederá a analizar el coeficiente de correlación de Pearson, lo que se hace es tomar una variable en dos periodos distintos. Se debe cumplir para que un instrumento de recolección de datos sea confiable el coeficiente de correlación ya mencionado debe ser superior al 0.70.

**Tabla 5. Coeficiente de correlación de Pearson para el logaritmo natural del ingreso por hora, periodo 2017 – 2018**

		In_ing_hora17	In_ing_hora18
In_ing_hora17	Correlación de Pearson	1	,794**
	Sig. (bilateral)		,000
	N	344	267
In_ing_hora18	Correlación de Pearson	,794**	1
	Sig. (bilateral)	,000	
	N	267	332

Nota: (\*\*) La correlación es significativa en el nivel 0,01 (bilateral).

La Tabla 5 nos muestra que el coeficiente de correlación de Pearson es de 0.794, por lo tanto, queda determinado que el instrumento es confiable.

**Tabla 6. Coeficiente de correlación de Pearson para la antigüedad en años, periodo 2017 – 2018**

		antigüedad_años17	antigüedad_años18
antigüedad_años17	Correlación de Pearson	1	,752**
	Sig. (bilateral)		,000
	N	666	599
antigüedad_años18	Correlación de Pearson	,752**	1
	Sig. (bilateral)	,000	
	N	599	650

Nota: (\*\*) La correlación es significativa en el nivel 0,01 (bilateral).

La Tabla 6 indica que el coeficiente de correlación de Pearson es de 0.752, por lo tanto, queda determinado que el instrumento es confiable.

**Tabla 7. Coeficiente de correlación de Pearson para las horas trabajadas, periodo 2017 – 2018**

		horas_trab17	horas_trab18
horas_trab17	Correlación de Pearson	1	,496**
	Sig. (bilateral)		,000
	N	666	599
horas_trab18	Correlación de Pearson	,496**	1
	Sig. (bilateral)	,000	
	N	599	650

Nota: (\*\*) La correlación es significativa en el nivel 0,01 (bilateral).

En la Tabla 7, tenemos un coeficiente de correlación por debajo del mínimo aceptable (0.496) para que instrumento de recolección de datos sea confiable. Sin embargo, esto podría cambiar significativamente si la encuesta fuera realizada en un periodo mensual o trimestral donde la condición laboral sea la misma, en este caso la encuesta fue realizada de manera anual. Por ende, esto significa que no depende de que el instrumento sea confiable o no, sino que va a depender de la condición laboral que ha tenido el periodo anterior y la condición actual del encuestado.

Por otro lado, si queremos ver la confiabilidad del instrumento en la variable edad, esperaríamos que la correlación sea perfecta, ya que la variable edad depende del tiempo y es inevitable para cualquier individuo.

Además, que se está trabajando con un panel de datos, esto significa que se ha hecho un seguimiento durante 5 años a estos individuos

**Tabla 8. Coeficiente de correlación de Pearson para la edad, periodo 2017 - 2018**

		edad17	edad18
edad17	Correlación de Pearson	1	1,000**
	Sig. (bilateral)		,000
	N	1071	1071
edad18	Correlación de Pearson	1,000**	1
	Sig. (bilateral)	,000	
	N	1071	1071

Nota: (\*\*) La correlación es significativa en el nivel 0,01 (bilateral).

Como era de esperar en la Tabla 8, el coeficiente de correlación de Pearson nos da un valor de 1.00. lo que significa una correlación perfecta y al mismo tiempo nos dice que es muy significativo.

#### 4.2. Análisis descriptivo

En este apartado se analizará como ha sido el comportamiento en el tiempo y en los grupos de análisis las variables de mayor relevancia en nuestro modelo.

En la Tabla 9, se presenta el tamaño y la distribución de la población de estudio.

**Tabla 9. Distribución de frecuencias de la población de estudio, periodo 2014 – 2018**

Año	Sexo		Total
	hombre	mujer	
2014	551	520	1,071
2015	552	519	1,071
2016	551	520	1,071
2017	551	520	1,071
2018	551	520	1,071
<b>Total</b>	<b>2,756</b>	<b>2,599</b>	<b>5,355</b>

La Tabla 9, nos indica que la muestra está conformada por dos grupos en un periodo de 5 años, esto significa que en todo este periodo los hombres representan el 51% y las mujeres un

49%. Además, el número total de individuos por año es de 1, 071 por lo que tenemos un acumulado de 5, 355 individuos durante el periodo de estudio.

Por otro lado, se tiene a la distribución de los ingresos mensuales para estos grupos. En el caso de los hombres sus ingresos han ido en aumento desde el 2014 al 2018. Sin embargo, este comportamiento no se da en el caso de las mujeres donde se encontró que para el periodo 2015 – 2016 sus ingresos cayeron, la caída no es significativamente grande, pero tiene sus implicancias en la brecha salarial. Es decir, el hecho de que el salario de las mujeres se estanque o en el peor de los casos disminuya (como en este caso) implica que la brecha va a ser mucho mayor ya que el salario de los hombres siempre va en aumento.

Además, si comparamos el promedio de este ingreso durante los 5 años se obtiene un salario para los hombres de 1,543.05 mientras que para las mujeres este promedio es de 1,183.64 lo que representa una diferencia de 359.41 soles. Sin embargo, las desviaciones de los ingresos respecto al promedio durante los 5 años son mayor para los hombres, siendo en el caso de estos el más alto en el año 2018 y en el caso de las mujeres en el año 2015.

El rango de los ingresos tanto para hombres como para mujeres fue en el año 2017, este estadístico al mediar la diferencia entre el valor máximo y el mínimo es lógico obtener el mayor rango en el año donde el ingreso fue mayor.

Por otro lado, si vemos el ingreso más alto en ambos grupos éste es para el año 2018. Con un ingreso promedio para los hombres de 1,681.79 y en el caso de las mujeres obtuvieron un ingreso 1,351.51 en ambos casos este ingreso fue superior a la media de los 5 años.

En el caso de la mediana este valor fue más alto en el caso de los hombres en el año 2017 y en el caso de las mujeres en el año 2018. En la cual, podemos decir que la mediana indica aquel 50% de la muestra está por debajo de ese ingreso y que la parte restante está por encima de ese nivel de ingreso determinado. Dicho de otro modo, en el caso de los hombres la mitad de ellos alcanzaron un ingreso de 1,200 soles y el otro 50% logro pasar este monto, en el caso de las mujeres la mitad de ellas gana 900 soles y el resto percibe un ingreso superior.

De manera detallada se presenta los estadísticos en la Tabla 10 para los hombres y en la Tabla 11 para las mujeres.

**Tabla 10. Medidas de tendencia central y de variabilidad para los hombres, periodo 2014 – 2018**

Summary for variables: ingreso\_mensual  
by categories of: year (Año)

year	mean	p50	sd	min	max	range
2014	1399.431	1100	1099.267	140	10000	9860
2015	1485.401	1200	1058.811	360	9000	8640
2016	1563.091	1200	1298.141	140	14000	13860
2017	1582.192	1257	1270.379	200	15000	14800
2018	1681.788	1250	1392.104	120	14500	14380
Total	1543.048	1200	1232.458	120	15000	14880

**Tabla 11. Medidas de tendencia central y de variabilidad para las mujeres, periodo 2014 – 2018**

Summary for variables: ingreso\_mensual  
by categories of: year (Año)

year	mean	p50	sd	min	max	range
2014	1063.367	800	931.4935	75	5700	5625
2015	1163.138	800	1189.569	300	8000	7700
2016	1131.224	850	898.3305	100	7000	6900
2017	1198.308	925	1003.056	240	8300	8060
2018	1351.509	1040	1018.847	140	6500	6360
Total	1183.635	900	1010.472	75	8300	8225

Hasta ahora hemos visto el ingreso promedio para ambos grupos. Por otro lado, nos resulta muy importante analizar dicho ingreso por nivel educativo y de qué manera se distribuye. En la Tabla 12, se detalla dicha distribución.

Esta tabla muestra el ingreso mensual promedio de los 5 años en la que se puede notar que a medida que el nivel educativo aumenta sus ingresos también aumenta, pero siempre existiendo una diferencia entre sexos. La diferencia más marcada a favor de los hombres se da en aquellas personas que tienen posgrado universitario.

El tema de las edades también es un factor muy importante para determinar la concentración del ingreso por rangos de edad.

En la Tabla 13, claramente podemos ver que el salario no se distribuye de la misma manera entre hombres y mujeres en los rangos de edades establecidos. Si bien es cierto que en todos los rangos de edad el salario de los hombres es siempre superior al de las mujeres se puede ver que el mayor ingreso para los hombres se encuentra en el rango de 26 a 40 años, en el caso de

las mujeres este cambia relativamente puesto que para las mujeres el rango de edades donde perciben mayores ingresos esta entre 41 y 65 años.

**Tabla 12. Distribución de los ingresos mensuales por nivel educativo**

Nivel de Educación	Sexo	
	hombre	mujer
Sin nivel	796.5	740.875
Primaria	966.647	714.1889
Secundaria	1372.21	884.6791
Sup no universitaria	1562.462	1161.033
Sup universitaria	2114.255	1851.891
Post - grado universitario	6651.278	2915.13
Total	1543.048	1183.635

Nota: Las distribuciones de estos ingresos han sido considerado para 5 años en su conjunto. Es decir, la tabla muestra los resultados de los ingresos promedio de este periodo.

**Tabla 13. Distribución de los ingresos mensuales por rangos de edad.**

Edad - Rangos	Sexo	
	hombre	mujer
De 14 a 25 años	1045.786	855.1625
Entre 26 y 40 año	1663.233	1196.283
Entre 41 y 65 año	1600.789	1283.979
Más de 65 años	1555.641	1178.333
Total	1543.048	1183.635

Nota: los niveles de ingreso mensual promedio de los 5 años.

A esto se le puede atribuir dos explicaciones, la primera es que a medida que pasa los años ambos sexos van adquiriendo conocimiento y con ello su rentabilidad aumenta. La segunda explicación va más por el tema la carga familiar, es decir, las mujeres entre el rango de edades entre 26 y 40 años se encuentran en una etapa fértil por lo que es posible que se tengan hijos y se dedique al cuidado de los hijos y además su participación en el mercado laboral baje y consigo sus ingresos sean inferiores. Pues, esto resulta evidente ya que a partir de los 41 años las féminas tienen un mayor nivel de ingreso.

Por otro lado, al analizar la lengua materna para ambos grupos y comparar sus ingresos, éstos siempre son mayores en el caso de que la lengua materna sea Castellano.

**Tabla 14. Distribución de los ingresos mensuales según lengua materna**

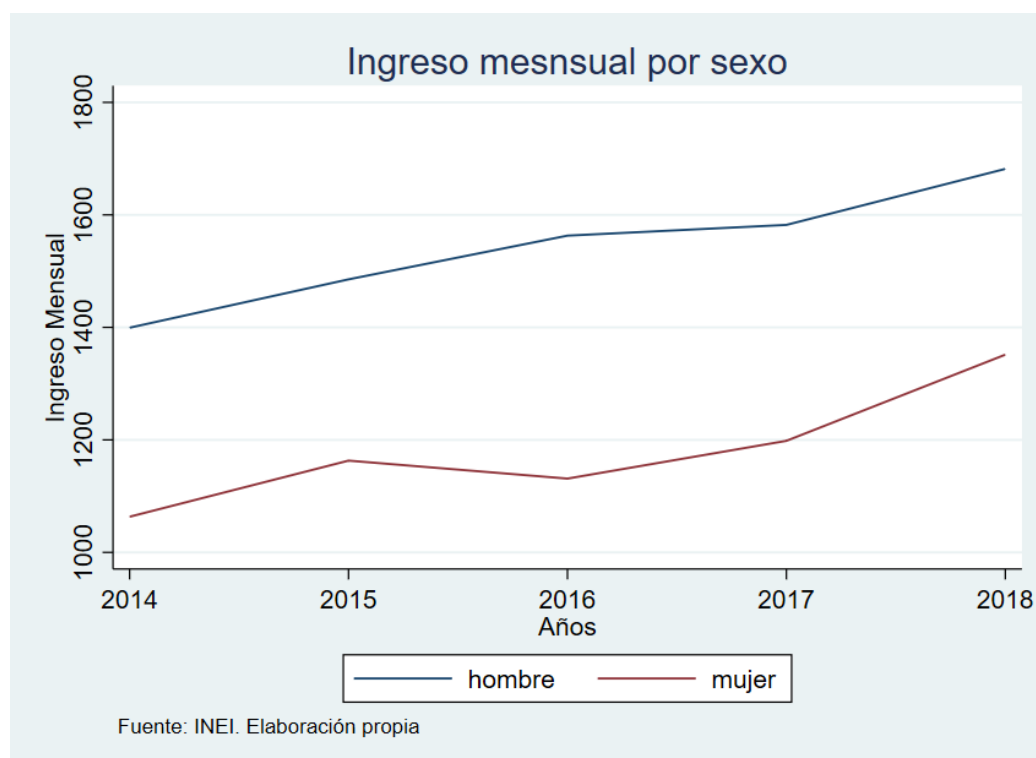
Lengua Materna	Sexo	
	hombre	mujer
Castellano	1555.195	1226.03
Otra lengua	1478.125	1166.8936
Total	1544.973	1183.635

*Nota:* los niveles de ingreso mensual promedio de los 5 años.

En el caso de los hombres el haber aprendido en su niñez cualquiera de estas lenguas el ingreso que percibe es superior a la Remuneración Mínima Vital (RMV) que equivale a 930.00 soles. En el caso de las mujeres solo se logra sobrepasar la RMV si en su niñez aprendió el Castellano.

El comportamiento del ingreso a lo largo del tiempo se puede ver en la Figura 4. En este grafico podemos encontrar que los ingresos ya mencionados tienen una tendencia positiva, pero siempre existe una brecha entre ambos sexos. Los factores que determinan la brecha salarial son el nivel educativo, la edad, la experiencia, etc.

**Figura 4. Ingreso mensual promedio para hombres y mujeres, periodo 2014 - 2018**



### 4.3. Análisis inferencial

#### *Coefficiente de correlación de Pearson:*

Dentro del análisis inferencial tenemos a los que más se adaptan para explicar nuestro problema. Una de ellas es el coeficiente de correlación de Pearson, éste conocido como coeficiente producto – momento analiza la relación entre dos variables y su condición es que estas variables tengan un nivel de medición por intervalos o de razón.

**Tabla 15. Coeficiente de correlación de Pearson: logaritmo del ingreso por hora con las horas trabajadas**

```
. pwcorr ln_ing_hora horas_trabajadas_op
```

	ln_ing~a	horas_~p
ln_ing_hora	1.0000	
horas_trab~p	-0.3204	1.0000

*Nota:* horas\_trab~p se refiere a las horas trabajadas.

En esta tabla podemos observar una correlación negativa, que nos indica que a medida que disminuye las horas trabajadas aumenta su nivel de ingresos del individuo. A esto se le conoce como efecto renta, esto nos indica que ante un incremento en el nivel de salario el individuo decide trabajar menos y consumir más ocio.

**Tabla 16. Coeficiente de correlación de Pearson: logaritmo del ingreso por hora con el nivel de educación**

```
. pwcorr ln_ing_hora nivel_estudios
```

	ln_ing~a	nivel_~s
ln_ing_hora	1.0000	
nivel_estu~s	0.4820	1.0000

*Nota:* el nivel\_estu~s es la variable nivel de educación.

Por otro lado, en la Tabla 16, se puede observar una correlación positiva, lo que indica que a medida que el individuo alcance un mayor nivel de estudios el ingreso también aumentara.

En la Tabla 17, se muestra la correlación de la variable dependiente con la edad.

La edad es otra de las variables que tiene una correlación positiva, indicando que a mayores años de edad se espera un mayor ingreso. Sin embargo, esto no siempre es así ya que hay momento en que la edad ya es muy alta y como consecuencia de estos el ingreso tiende a bajar.

**Tabla 17. Coeficiente de correlación de Pearson: logaritmo del ingreso por hora con la edad**

```
. pwcorr ln_ing_hora edad
```

	ln_ing~a	edad
ln_ing_hora	1.0000	
edad	0.1457	1.0000

El tamaño de la empresa es otra variable que se correlaciona de manera positiva con nuestra variable dependiente, esto indica que el individuo que labore en una empresa de mayor tamaño también percibe un ingreso mayor. Esta correlación se expresa en la Tabla 18.

**Tabla 18. Coeficiente de correlación de Pearson: logaritmo del ingreso por hora con el tamaño de la empresa**

```
. pwcorr ln_ing_hora trabajo_laboraron
```

	ln_ing~a	trabaj~n
ln_ing_hora	1.0000	
trabajo_la~n	0.4391	1.0000

*Nota:* trabajo\_la~n se refiere al tamaño de la empresa expresado por el número de trabajadores que laboran en ella.

**Tabla 19. Coeficiente de correlación de Pearson: logaritmo del ingreso por hora con la antigüedad en la empresa**

```
. pwcorr ln_ing_hora tiempo_op_años
```

	ln_ing~a	tiemp~os
ln_ing_hora	1.0000	
tiempo_op~os	0.2782	1.0000

*Nota:* tiempo\_op~os es la antigüedad de la ocupación que desempeña medida en años, para la interpretación de resultados esta variable es explicada como la experiencia del individuo.

Otra de las variables que tiene una correlación positiva es tiempo que el individuo lleva trabajando en esa ocupación o la experiencia. Entonces, a mayor experiencia, mayores ingresos.

### ***Test de varianza:***

Ahora, vamos a determinar si las varianzas poblacionales de nuestro grupo que está conformado por hombres y mujeres son iguales, a esto también se le llama homogeneidad de varianza o homocedasticidad. Este planteamiento da mayor validez y consistencia al momento de la estimación e interpretación de resultados.

Para ello, vamos a realizar un test de varianza que está realizada por la prueba de Levene que pone a prueba la hipótesis nula que las varianzas son iguales.

En la Tabla 20, podemos observar la hipótesis alternativa es significativa al 95% y, por lo tanto, se rechaza la hipótesis nula de que las varianzas son iguales. Además, con esto afirmamos que la desviación observada es estadísticamente diferente entre ambos grupos.

**Tabla 20. Test de varianza**

```
. sdtest ln_ing_hora, by(sexo)
```

Variance ratio test

Group	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf. Interval]	
hombre	1,150	1.847711	.0180385	.6117163	1.812319	1.883103
mujer	509	1.567163	.0308432	.6958537	1.506567	1.627758
combined	1,659	1.761636	.0159948	.6514819	1.730264	1.793008

ratio = sd(hombre) / sd(mujer) f = 0.7728  
 Ho: ratio = 1 degrees of freedom = 1149, 508

Ha: ratio < 1 Pr(F < f) = 0.0002	Ha: ratio != 1 2*Pr(F < f) = 0.0005	Ha: ratio > 1 Pr(F > f) = 0.9998
-------------------------------------	--	-------------------------------------

### ***Test de medias:***

Las pruebas de dos muestras nos permiten contrastar el hecho de que las medias de dos grupos son iguales utilizando para ello las medias de dos muestras aleatorias extraídas de esa población. La hipótesis nula que plantea esta prueba es que el promedio de los ingresos (en términos logarítmicos) sean idénticos entre hombres y mujeres.

Tal y como se puede ver en la Tabla 21, donde el ingreso de los hombres es mayor al de las mujeres, lo que tendremos que analizar ahora es que esta diferencia sea estadísticamente diferente.



ello, se estima una regresión general (tomando en cuenta los dos grupos), los resultados se expresan en la Tabla 23.

Siguiendo con el mismo nivel de significancia del p-value de 5% o 0.05, tenemos a las variables que no son significativas en el modelo, estas variables son la situacion\_sentimental (situación familiar), la lengua, las atenciones\_parto (atenciones de parto), las limitaciones para moverse, la antigüedad en el distrito y el pueblo indígena.

El modelo general tiene un R2 global (overall) de 41.76% esto porcentaje nos indica la significancia del modelo, en otras palabras, el R2 se refiere al grado en que explican las variables independientes a la variable dependiente, además, este nos indica que el modelo con datos de panel explica el efecto que existe sobre el tiempo e individuos. El efecto entre individuos (between) es de 40.93% y el efecto en el tiempo (within) es de 15.82%.

**Tabla 23. Regresión panel para ambos grupos**

Random-effects GLS regression		Number of obs = 1,624				
Group variable: ID		Number of groups = 473				
R-sq:		obs per group:				
within = 0.1582		min = 1				
between = 0.4093		avg = 3.4				
overall = 0.4176		max = 5				
corr(u_i, x) = 0 (assumed)		wald chi2(12) = 537.49				
		Prob > chi2 = 0.0000				
ln_ing_hora	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
situacion_familiar	-.045307	.0245317	-1.85	0.065	-.0933882	.0027742
edad	.04156	.0084824	4.90	0.000	.0249347	.0581852
edad2	-.0004107	.0000972	-4.23	0.000	-.0006012	-.0002202
nivel_estudios	.1665859	.0169712	9.82	0.000	.1333229	.1998488
lengua	.0243738	.0418513	0.58	0.560	-.0576534	.1064009
atenciones_parto	.0343404	.196568	0.17	0.861	-.3509258	.4196066
limitaciones	-.0552889	.1339864	-0.41	0.680	-.3178974	.2073196
antiguedad_distrito	-.0123014	.0530582	-0.23	0.817	-.1162935	.0916907
horas_trabajadas_op	-.0132416	.0009644	-13.73	0.000	-.0151318	-.0113514
tiempo_op_años	.0064834	.0017198	3.77	0.000	.0031128	.0098541
pueblo_indigena	-.011858	.014397	-0.82	0.410	-.0400755	.0163596
trabajo_laboraron	.0762018	.0081525	9.35	0.000	.0602232	.0921804
_cons	.7530466	.5298388	1.42	0.155	-.2854184	1.791512
sigma_u	.42185612					
sigma_e	.29021541					
rho	.67876057	(fraction of variance due to u_i)				

Por lo tanto, al realizar las estimaciones por separado considerando solamente las variables significativas para el modelo se obtuvo los siguientes resultados, ver Tabla 24.

En el caso de los hombres el R2 global es de 37.10%, el efecto entre individuos es de 36.61% y la variabilidad o efecto en el tiempo es de 13.98%.

Por otro lado, se tiene que todos los coeficientes son significativos teniendo en cuenta un nivel de confianza del 95% y su nivel de significancia del 5% o 0.05 y cumpliendo con el objetivo principal de la investigación se tiene que los factores determinantes de la brecha salarial entre hombres y mujeres en el mercado laboral peruano: 2014 – 2018 son la edad, la edad elevada al cuadrado, el nivel de estudios, las horas trabajadas, el tiempo en su ocupación principal y el tamaño de la empresa.

**Tabla 24. Regresión panel para los hombres**

Random-effects GLS regression	Number of obs	=	1,115
Group variable: ID	Number of groups	=	313
R-sq:	obs per group:		
within = 0.1398	min =		1
between = 0.3661	avg =		3.6
overall = 0.3710	max =		5
corr(u_i, x) = 0 (assumed)	wald chi2(6)	=	305.81
	Prob > chi2	=	0.0000

ln_ing_hora	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
edad	.0444792	.0092454	4.81	0.000	.0263586 .0625999
edad2	-.0004387	.0001064	-4.12	0.000	-.0006472 -.0002301
nivel_estudios	.1413564	.0206486	6.85	0.000	.1008859 .1818269
horas_trabajadas_op	-.0128086	.0011481	-11.16	0.000	-.0150588 -.0105583
tiempo_op_años	.0044352	.0018861	2.35	0.019	.0007385 .0081319
trabajo_laboraron	.0702472	.0096806	7.26	0.000	.0512735 .0892209
_cons	.753341	.2119119	3.55	0.000	.3380014 1.168681
sigma_u	.40550547				
sigma_e	.29369215				
rho	.65592901	(fraction of variance due to u_i)			

La regresión panel para las mujeres se encuentra plasmada en la Tabla 25.

A diferencia de los hombres, el modelo estimado para las mujeres tiene un R2 global más elevado, este modelo explica el efecto del tiempo e individuos en un 51.98% de los cuales se desprende el efecto en el tiempo de 21.49% y el efecto entre individuos de 49.86%. Cabe precisar que la variabilidad entre individuos es más alta en ambos grupos.

Sin embargo, para poder hacer las interpretaciones de los coeficientes es necesario saber la descomposición de estos. Por lo que se presentan los resultados en una misma tabla, en la cual podemos en la Tabla 26.

**Tabla 25. Regresión panel para las mujeres**

Random-effects GLS regression	Number of obs =	509
Group variable: ID	Number of groups =	161
<b>R-sq:</b>	Obs per group:	
within = 0.2149	min =	1
between = 0.4986	avg =	3.2
overall = 0.5198	max =	5
corr(u_i, x) = 0 (assumed)	wald chi2(6) =	252.75
	Prob > chi2 =	0.0000

ln_ing_hora	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
edad	.0596815	.0156437	3.82	0.000	.0290204 .0903427
edad2	-.0006386	.0001925	-3.32	0.001	-.0010158 -.0002613
nivel_estudios	.1920661	.0275261	6.98	0.000	.1381159 .2460163
horas_trabajadas_op	-.014119	.0017061	-8.28	0.000	-.0174629 -.0107751
tiempo_op_años	.0123028	.0039298	3.13	0.002	.0046006 .020005
trabajo_laboraron	.0825491	.0144093	5.73	0.000	.0543074 .1107908
_cons	.023095	.3325611	0.07	0.945	-.6287128 .6749028
sigma_u	.41267089				
sigma_e	.27833952				
rho	.68731952	(fraction of variance due to u_i)			

**Tabla 26. Comparación de las estimaciones de los grupos**

variable	Hombres	Mujeres
Edad	.04327617	.0592476
Edad al Cuadrado	-.00042694	-.00063553
Nivel de Educación		
Primaria	.23477437	-.01198495
Secundaria	.41452459	.2322259
Sup no universitaria	.49244369	.48060613
Sup universitaria	.66675013	.62315114
Post - grado universitario	.99509761	.77745018
# de Horas Trabajadas	-.01282926	-.01383207
Antigüedad - Años	.00466506	.0129894
Tamaño de la Empresa		
de 21 a 50 personas	.18240205	.1354206
de 51 a 100 personas	.19287952	.33800682
de 101 a 500 personas	.28693112	.3796192
más de 500 personas	.29316518	.30563347
Constant	.8559725	.42694498

**Edad.** La edad en la determinación de los ingresos viene representada con una curva cóncava, si bien es cierto que a medida que aumente la edad el ingreso también aumenta, este

cambia a una pendiente negativa cuando se llega a un punto máximo. Es decir, ese punto máximo indica que desde allí el ingreso empieza a decaer.

Para ello, vamos a determinar cuál es el máximo de horas trabajadas para ambos sexos.

De esta manera, valiéndonos de la teoría económica y matemática este cálculo se realiza de la siguiente manera:  $\frac{d \ln(\text{ing\_hora})}{d(\text{edad})} = \beta_1 + 2\beta_2 \text{edad}$ , para el cálculo manual esto sería:

$x^* = \left| \frac{\beta_1}{2\beta_2} \right|$ , al realizar este cálculo con los estimadores obtenidos encontramos un punto

máximo para los hombres de 50.68 años. Es decir, que a partir de ese punto los ingresos de los hombres empieza a ser decreciente.

Sin embargo, para las mujeres este punto máximo es de 46.61 años. Es decir, las mujeres empiezan a percibir menos ingresos aproximadamente 4 años antes que los hombres.

**Figura 5. Punto máximo de edad: Hombres, Mujeres y promedio**



En la Figura 5, se puede ver la línea vertical de cada gráfico que indica que desde ese punto los ingresos empiezan a caer. Si se considera dentro del cálculo a ambos sexos se obtiene el punto máximo de 51.92 años.

**Nivel educativo.** Esta variable muestra una relación positiva con el ingreso. Es decir, a medida que aumenta el nivel educativo también aumentara su ingreso. Pero, lo que nos interesa saber es en cuanto aumenta el ingreso cuando aumenta el nivel educativo, la explicación de los coeficientes obtenidos de las regresiones se contempla en la Tabla 27.

En base a esto, tomando como base a las personas que no tienen ningún nivel educativo (Sin nivel), tenemos:

*Tabla 27. Interpretación del nivel educativo para hombres y mujeres*

Nivel Educativo	Hombres	Mujeres
Primaria	Ganan 23.48% más que las personas “Sin nivel”	Ganan 1.19% menos que las personas “Sin nivel”
Secundaria	Ganan 41.52% más que las personas “Sin nivel”	Ganan 23.22% más que las personas “Sin nivel”
Superior no universitaria	Ganan 49.24% más que las personas “Sin nivel”	Ganan 48.06% más que las personas “Sin nivel”
Superior universitaria	Ganan 66.67% más que las personas “Sin nivel”	Ganan 62.32% más que las personas “Sin nivel”
Post – grado universitario	Ganan 99.51% más que las personas “Sin nivel”	Ganan 77.75% más que las personas “Sin nivel”

**Horas trabajadas.** Las horas trabajadas tienen un efecto negativo con el ingreso, esto se explica la siguiente manera, a medida que disminuye las horas trabajadas aumenta su nivel de ingresos del individuo. A esto se le conoce como efecto renta, esto nos indica que ante un incremento en el nivel de salario el individuo decide trabajar menos y consumir más ocio.

**Antigüedad en la empresa.** Esta variable en el caso de los hombres nos indica que un incremento de la experiencia en un año genera un incremento en el ingreso del 0.46%. En el caso de las mujeres, un año adicional de experiencia hace que el ingreso aumente en 1.29%. Es decir, que la experiencia de las mujeres es más valorada que la de los hombres.

**Tamaño de la empresa.** El tamaño de la empresa también tiene una relación positiva con el nivel de ingresos. Si tomamos como base a las personas que trabajan en empresas de hasta 20 personas, tenemos los coeficientes, y su interpretación en la Tabla 28.

**Tabla 28. Interpretación del tamaño de empresa**

<b>Tamaño de Empresa</b>	<b>Hombres</b>	<b>Mujeres</b>
De 21 a 50 personas	Ganan 18.24% más que aquellos que trabajan en empresas de hasta 20 personas	Ganan 13.54% más que aquellas que trabajan en empresas de hasta 20 personas
De 51 a 100 personas	Ganan 19.29% más que aquellos que trabajan en empresas de hasta 20 personas	Ganan 33.80% más que aquellas que trabajan en empresas de hasta 20 personas
De 101 a 500 personas	Ganan 28.69% más que aquellos que trabajan en empresas de hasta 20 personas	Ganan 37.96% más que aquellas que trabajan en empresas de hasta 20 personas
Más de 500 personas	Ganan 29.32% más que aquellos que trabajan en empresas de hasta 20 personas	Ganan 30.56% más que aquellas que trabajan en empresas de hasta 20 personas

A partir de las empresas con más de 50 personas el coeficiente para las mujeres es más alto que el de los hombres. Por otro lado, en caso de los hombres ganan un mayor ingreso cuando el tamaño de la empresa es superior a 500 personas. Caso contrario ocurre con las mujeres, ellas perciben un mayor ingreso en empresas que tienen menos de 500 trabajadores.

Sin embargo, hasta estas alturas hemos encontrado y dado explicación a los factores determinantes de la brecha salarial y su impacto en los ingresos. Si bien es cierto que hemos visto una diferencia en el nivel de ingresos entre hombres y mujeres (ver Tabla 22) no sabemos con claridad que proporción de ésta viene explicada por la acumulación de capital humano.

Para ello, tenemos que proceder a la descomposición de la brecha salarial la cual está detallada en la Tabla 29. Es así, que el rectángulo de color rojo de la Tabla 29 indica que la media geométrica del logaritmo natural del ingreso por hora para los hombres es de 1.825 y de 1.567 para las mujeres, la diferencia entre sexos que se está señalada con una elipse de color verde en la tabla indica que el diferencial es de 0.258 a favor de los hombres.

En respuesta al primer objetivo específico se tiene que la brecha salarial de género a favor de los hombres es de 25.76%, esta es la brecha general, por lo que las dotaciones de los individuos dentro de la tabla están representadas como “endowments”, esto indica cuánto debería ser el incremento medio en el salario de las mujeres si tuvieran las mismas características individuales, es decir, las mismas dotaciones de capital humano que los hombres (nos referimos únicamente a los valores en promedio). Siendo así, el logaritmo de su ingreso

debería incrementar en 0.0033 en términos porcentuales debería incrementar 0.33%. Sin embargo, un aspecto que llama la atención de este resultado es que este componente observable no es significativo ya que el p-value es de 0.905 lo que indica que es mayor al 0.05, por lo que se dice que las características individuales no son significativas para explicar la diferencia salarial.

**Tabla 29. Descomposición de la brecha salarial de género en términos logarítmicos, 2014 – 2018**

ln_ing_hora	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
<b>Blinder-Oaxaca decomposition</b>						
				Number of obs	=	1,624
				Model	=	linear
Group 1: sexo = 1				N of obs 1	=	1115
Group 2: sexo = 2				N of obs 2	=	509
<b>overall</b>						
group_1	1.824775	.0180903	100.87	0.000	1.789319	1.860232
group_2	1.567163	.0309279	50.67	0.000	1.506545	1.62778
difference	.2576129	.0358301	7.19	0.000	.1873872	.3278386
endowments	.0032842	.0274476	0.12	0.905	-.0505122	.0570806
coefficients	.2760519	.0262361	10.52	0.000	.2246301	.3274737
interaction	-.0217232	.0109802	-1.98	0.048	-.043244	-.0002023

*Nota:* el group\_1 se refiere a los hombres y el group\_2 a las mujeres.

En el caso de los “coefficients” sucede lo contrario. Es decir, explica el cambio del nivel de ingresos de los hombres si se estimara tomando los coeficientes de las mujeres. Si dichas dotaciones (características individuales) de las mujeres se remunerarán a los mismos precios que el de los hombres, el logaritmo natural del ingreso de las mujeres incrementaría en 0.2761 lo que se traduce a un 27.61%, está es la proporción que explica el componente observable a la brecha salarial de género (o lo que es lo mismo las características propias del individuo).

Sin embargo, cuando se considera el efecto conjunto (interaction) de estos coeficientes sobre las dotaciones de capital humano de los hombres, el logaritmo de su ingreso por hora debería de disminuir en 0.0217 esto hace que se produzca un impacto neto positivo de 25.43%, este resultado se obtiene de la diferencia entre (0.2761 – 0.0217). Este último componente de la tabla es atribuible a la diferencia de los salarios no explicado, en otras palabras, nos referimos al componente discriminatorio de la brecha salarial y esté influye en 25.43% en la determinación de está.

Es decir, éste explica el impacto originado en que las diferencias entre las características propias de los individuos y en los coeficientes existen de manera sincrónica. Entonces, basándonos en estos resultados podemos decir que este efecto simultaneo es la heterogeneidad

no observada en el mercado laboral peruano, esto es a lo que se le puede atribuir la posible discriminación entre hombres y mujeres.

En contraste a lo anterior se puede apreciar que las dotaciones explican las diferencias salariales entre sexos solo un 1.09%, mientras tanto los coeficientes o lo que se le denomina efecto precio y que también está determinado por las características del individuo explican al diferencial en 91.69%. Por último, se tiene a la interacción que básicamente se refiere al componente discriminatorio llega a explicar solo un 7.22%.

Si bien es cierto que la brecha salarial de género en nuestro país ha decrecido significativamente en los últimos años aún es significativa. Según el (Instituto Nacional de Estadística e Informática , 2018) encontró que la brecha salarial para nuestro país en el año 2017 fue de 29.30%. En nuestra investigación se encontró una brecha del 25.76% claramente viendo que está a disminuido en ese periodo.

Además, se encontró que la parte explicada que determina la brecha salarial de género, es decir, las características propias de los individuos explican un incremento en el salario de un 0.33% y el efecto precio un 27.61%. Mientras tanto, la parte no explicada de la brecha salarial que se le atribuye a la discriminación influye con 25.43% en la determinación de la brecha salarial.

En contraste a lo anterior y de acuerdo con los objetivos específicos planteados en la investigación se encontró que las características propias del individuo explican a la brecha salarial en un 92.78% (siendo este resultado de la suma del 0.33% + 91.69%) mientras que el factor discriminatorio llega a explicar solo un 7.22%.

En efecto, si nos comparamos con un país como España con una brecha de género del 12.20% al 2018 según un estudio elaborado por las internacionales (CEOE & PwC, 2018) la diferencia es aún más abismal.

Así mismo, en un estudio para la economía chilena en el año 2003 por (Fuentes, et al. 2005) muestra que la brecha era de un 27.5%. Lo que podemos rescatar de este estudio es que a pesar de a ver pasado aproximadamente 17 años la brecha es similar a la nuestra. Esto nos quiere decir que la economía chilena tiene una brecha salarial de género inferior a la nuestra.

Finalmente, tenemos los seis métodos de descomposición de Oaxaca – Choe, donde viene representado de la siguiente manera en la Tabla 30.

Dado que el objetivo principal de la tesis no fue explicar a detalle la descomposición de esta brecha se hace solamente una comparación de la metodología empleada en Alemania en el año 2016 donde se encontró una brecha en este país de 27.7%, es preciso mencionar que el periodo de estudio para este país fue de 1986 – 2011 es por ello por lo que la brecha encontrada es aún

elevada a comparación de los países de Europa en la actualidad. Sin embargo, en nuestro país se encontró una brecha salarial de género del 25.76% y estamos por encima de la media de lo que esta Latinoamérica con 20% aproximadamente.

**Tabla 30: Descomposición de la brecha salarial entre hombres y mujeres**

	Ei		Ui		Si		Brecha Salarial
Método 1	0.09144800	35.50%	0.16615200	64.50%			0.2576
Método 2	0.11965520	46.45%	0.13794480	53.55%			0.2576
Método 3	0.09144800	35.50%	0.13804784	53.59%	0.02810416	10.91%	0.2576
Método 4	-0.02369920	-9.20%	0.13132448	50.98%	0.14997472	58.22%	0.2576
Método 5	0.02782080	10.80%	0.22977920	89.20%			0.2576
Método 6	0.00600208	2.33%	0.23155664	89.89%	0.02004128	7.78%	0.2576

### **Sesión de discusión**

El objetivo principal de la tesis fue identificar los factores determinantes de la brecha salarial entre hombres y mujeres en el mercado laboral peruano para el periodo 2014 – 2018, haciendo la corrida del modelo econométrico y aplicar la metodología señalada se encontró que estos factores que lo determinan son: la edad, la edad al cuadrado, el nivel educativo, las horas trabajadas, el tiempo en su ocupación principal y el tamaño de la empresa.

En cuanto al componente explicado de la brecha salarial, tenemos a la edad como primer factor determinante. Esta variable muestra una correlación positiva con el nivel de ingresos, sin embargo, está es una variable que tiene una forma cuadrática si es que la elevamos al cuadrado, esto quiere decir que si bien es cierto que está tiene una relación positiva con el nivel de ingresos en algún momento esta relación cambia y se vuelve negativa así como lo muestra la edad elevada al cuadrado en la regresión de la Tabla 26 donde se puede ver que para ambos grupos su signo es negativo. En la mayoría de los trabajos de investigación se llega hasta este punto (demostrar la relación positiva de la edad con el nivel de ingresos) sin embargo, en esta tesis se logró calcular cuál es ese punto máximo (de la edad) en el que la relación pasa a ser inversa, ver Figura 5.

En cuanto al nivel educativo, tenemos que para ambos grupos el coeficiente es positivo lo que significa que a mayor años de educación mayor es el nivel de ingresos que se percibe, concordamos con las investigaciones de (Cortés Aguilar & Flórez Vera, 2015) en el caso del mercado colombiano en la provincia de Santander para ser preciso, en el caso peruano tenemos a la investigación de (Cahuas, 2019). Es por ello, en esta última encontramos que el coeficiente para los hombres es de 0.044 y para las mujeres de 0.033, esto significa que por un año adicional de educación hace que aumente el nivel de ingreso en 4.4 y 3.3% para los hombres y mujeres respectivamente. Sin embargo, en esta tesis se encontró que en el caso de los hombres un año

de educación adicional incrementa su ingreso en 14.14% mientras que para las féminas este coeficiente es mayor y hace que ganen 19.20% más. A modo de comparación se tiene que en el trabajo elaborado (Cahuas, 2019) se encontró que un año de educación adicional tenía un impacto mayor en la determinación del nivel de ingresos de los hombres, caso contrario ocurre en nuestro hallazgo ya que se encentra que un año adicional de educación incrementa en mayor proporción el nivel de ingresos de las mujeres.

Tomando nuevamente como referencia el trabajo de (Cahuas, 2019) se tiene que el tamaño de la empresa para los hombres es negativo. Es decir, mientras mayor sea el tamaño de la empresa los hombres tendrán ingresos más bajos excepto en las empresas donde el número de trabajadores sobrepasa los 500. Caso contrario ocurre con las mujeres, donde se puede apreciar que en todos los casos el nivel de ingresos aumenta, para las mujeres mientras más grande sea la empresa mayor será el nivel de ingresos que se percibirá. En nuestro caso se comparte igual que lo último para ambos grupos. Es decir, mientras mayor sea el tamaño de la empresa mayor será el nivel de ingresos tanto de hombres como el de las mujeres. No obstante, es esencial precisar que el salario de las mujeres aumenta en mayor proporción a la de los hombres según el tamaño de la empresa.

Ahora bien, pasamos al componente discriminatorio de la brecha salarial y encontramos en un estudio elaborado por (Cortés Aguilar & Flórez Vera, 2015) que sostiene que el factor discriminatorio explica entre un 25% y 30% las diferencias salariales entre géneros en el mercado laboral colombiano. A su vez, en esta tesis se encontró que el factor discriminatorio explica tan solo el 7.22% de la brecha salarial de género. Esto quiere decir que la parte de la brecha salarial explicada por la discriminación en el mercado laboral peruano es pequeña pero no se puede decir que no es importante. A pesar de eso, si tomamos la brecha salarial de género en nuestro país que resulta 25.76% aún es alta a comparación de la región que tiene una brecha de 20%, según el (World Economic Forum, 2018).

A pesar de haber cumplido con el objetivo general y así también con los objetivos específicos el tema de la brecha salarial es muy amplio y tiene enfoques diversos, en particular esta tesis consistía en encontrar cuáles eran los factores determinantes de la brecha salarial de género y saber qué proporción tanto de la parte observable y la parte inobservable explican a la brecha salarial. Es por ello, que resulta pertinente mencionar que la brecha salarial no simplemente está presente entre hombres y mujeres, sino que también está presente en muchos grupos del mercado laboral. Por ende, se busca que futuras investigaciones continúen con los hallazgos de estas brechas y los factores que la determinan.

La única limitación que se tuvo fue con la base de datos al momento de incluir la experiencia del trabajador. Esta variable no se encuentra descrita tal cual en la base de datos de la ENAHO panel, la variable más próxima y que puede capturar el mismo efecto en el mercado laboral peruano fue la antigüedad del trabajador en el puesto de trabajo.

## V. Conclusiones

La tesis tuvo como objetivo general identificar los factores determinantes de la brecha salarial entre hombres y mujeres en el mercado laboral peruano:2014 – 2018. De tal objetivo obtuvimos mediante la aplicación de la metodología de Oaxaca – Choe que dichos determinantes son la edad, la edad al cuadrado, el nivel educativo, las horas trabajadas, el tiempo en su ocupación principal y el tamaño de la empresa.

De esta manera se concluye que la brecha salarial de género en nuestro país ha venido decreciendo, pero es aún elevada a comparación del resto de países de la región y ni que decir de los países desarrollados, la brecha salarial es un problema socioeconómico que genera un alto costo social a la economía mundial y nosotros no somos ajenos a ello ya que estamos con una brecha salarial de género de 25.76% sobre la media de Latinoamérica que es del 20%.

Teniendo en cuenta el nivel educativo, la antigüedad en su ocupación principal y el tamaño de la empresa que son los factores que determinan la brecha salarial y a su vez tienen una relación positiva en la reducción de esta es de gran relevancia concluir que a pesar de ello esta brecha no se puede eliminar simplemente mejorando estos factores, es cierto que podemos esperar niveles mínimos de ésta, pero resulta imposible poder eliminarla es por ello por lo que necesitamos la intervención del gobierno mediante políticas públicas alineadas a la igualdad salarial entre hombres y mujeres.

Por otro lado, se concluye que las características propias de los individuos explican un 27.61% en la determinación de los ingresos, así mismo, este componente representa un 92.78% de la brecha salarial. Sin embargo, el factor discriminatorio llega a determinar el 25.43% el nivel de ingresos de los individuos y explica la brecha salarial un 7.22%.

## **VI. Recomendaciones**

Siguiendo con esta línea de investigación se recomienda seguir con estudios futuros de esta índole ya que nos permiten identificar los factores que determinan el nivel de ingresos y genera las diferencias salariales. Además, con esto se sabrá en que aspectos debemos mejorar como país.

En cuanto a la metodología en esta tesis se utilizó la más reciente que es la de Oaxaca – Choe que divide a la brecha salarial en seis componentes permitiendo explicar a más detalle la brecha salarial de género, pero de igual manera se recomienda seguir creando nuevas metodologías que permitan descomponer la brecha salarial.

La brecha salarial de género convive en nuestra sociedad y se ha llegado a la conclusión que, aunque mejoremos al máximo en los factores que ayudan a reducir esta brecha esta es imposible erradicarla, para ello se recomienda la intervención del gobierno mediante políticas públicas dirigidas a la igualdad de salarios propiciando mejores condiciones laborales para las mujeres peruanas y de esta manera contribuir con el desarrollo empresarial generando un impacto positivo en nuestra economía.

## VII. Lista De Referencias

- Alfaro, D., & Guerrero, E. (2013). Brechas de género en el ingreso: Una mirada más allá de la media en el sector agropecuario. *Consortio de investigación Económica y Social (CIES)*.
- Anghel, B., Conde, J. I., & Marra, I. (2018). Brechas salariales de género en España. *Federación de Estudios de Economía Aplicada*, 30.
- Barbosa, W. J., & Castañeda, J. B. (2017). *Determinantes de las diferencias salariales y laborales por género, un análisis econométrico para Colombia entre los años 2005 - 2014*. Bogotá D.C.: Universidad de la Salle .
- Benítez, D., & Espinoza, B. (2018). *Discriminación salarial por género en el sector formal en Ecuador usando registros administrativos*. Quito: Instituto Nacional de Estadística y Censos - INEC.
- Bernal, C. (2010). *Metodología de la Investigación*. Colombia: Prentice Hall.
- Brindusa, A., J. Ignacio, C., & Marra de Artúñano, I. (2018). *Brechas salariales de género en España*. Fundación de Estudios de Economía Aplicada .
- Cahuas, P. O. (2019). *Educación y brecha salarial de género en el Perú*. Lima: PUCP.
- CEOE & PwC. (2018). *Análisis de la brecha salarial de género en España*. PwC.
- Chamberlain, A., Zhao, D., & Stansell, A. (2019). Progress on the Gender Pay Gap: 2019. *Glosdoor, Economic Research*, 79.
- Comisión de Igualdad y Diversidad de CEOE. (Marzo de 2019). Análisis de la brecha salarial de género en España. (S. PricewaterhouseCoopers Asesores de Negocios, Ed.) *Confederación Española de Organizaciones Empresariales*, 70.
- Comisión Europea. (2014). Cómo combatir la brecha salarial entre hombres y mujeres en la Unión Europea. *Unión Europea*, 28.
- Cortés Aguilar, A., & Flórez Vera, M. A. (2015). *Diferencias salariales por género en el departamento de Santander, Colombia*. Santander: CENES.
- Cozarla, M. Y. (2016). Discriminación y brecha salarial por género en el Perú . *Universidad Nacional del Altiplano*, 75.
- Díaz, E. (2015). *La desigualdad salarial entre hombres y mujeres*. Santiago: Departamento de Estudios de la Dirección del Trabajo.
- Fuentes, J., Palma, A., & Montero, R. (2005). Discriminación Salarial en Chile: Una mirada global. *ResearchGate*, 26.
- Galassi, G., & Andrada, M. (2009). La relación entre educación e ingresos: ecuaciones de Mincer por regiones geográficas de Argentina. *X Jornadas Argentinas de Estudios de Población. Asociación de Estudios de Población de la Argentina, San Fernando del Valle de Catamarca*, 29.
- Hernández, R. (2014). *Metodología de la Investigación*. McGRAW-HILL / INTERAMERICANA EDITORES, S.A. DE C.V.
- Hernandez, R., Fernandez, C., & Baptista, M. (2010). *Metodología de la investigación*. Mexico: McGraw-Hil.
- Instituto Nacional de Estadística e Informática . (2018). *Perú: Brechas de Género 2018*. Lima: Instituto Nacional de Estadística e Informática .
- Jung, S., Choe, C., & Oaxaca, R. L. (2016). Gender Wage Gaps and Risky vs. Secure Employment: An Experimental Analysis. *The Institute for the Study of Labor (IZA)*.
- McConnell, C. R., Brue, S. L., & Macpherson, D. A. (2003). *Economía Laboral* (Vol. 6). Aravaca, Madrid , España : McGraw-Hill/Interamericana de España, S.A.U.
- Ministerio de Trabajo y Promoción del Empleo. (2016). La mujer en el mercado laboral peruano. *Ministerio de Trabajo y Promoción del Empleo*, 85.

- MTPE, M. d. (2017). Brecha Salarial por Género en el Sector Privado Formal en el Perú 2004-2016. *Ministerio de Trabajo y Promoción del Empleo*, 106.
- Oaxaca, R. L., & Choe, C. (2016). Wage Decompositions Using Panel Data Sample Selection Correction. *IZA*, 26.
- OIT. (2018). La brecha salarial en América Latina. *Organización Internacional de Trabajo*, 20.
- Organización Internacional de Trabajo. (2016). Las mujeres en el trabajo. *Tendencias*, 149.
- Organización Internacional de Trabajo. (2019). *Informe mundial sobre salarios*. Banco Mundial.
- Rabanal, J. E. (2001). *Dispersión salarial, capital humano y segmentación laboral en Lima*. Lima: Consorcio de Investigación Económica y Social; Pontificia Universidad Católica del Perú.
- Rodríguez Pérez, R. E., & Limas Hernández, M. (2017). El análisis de las diferencias salariales y discriminación por género por áreas profesionales en México, abordado desde un enfoque regional, 2015. *Revista de Alimentación Contemporánea y Desarrollo Regional*, 30.
- Salkind, N. (1998). *Métodos de Investigación* (Vol. III). México: Prentice-Hall.
- T. Denning, J., Jacob, B., Lefgren, L., & vom Lehn, C. (2019). *The return to hours worked within and across occupations: implications for the gender wage gap*. Massachusetts Avenue: National Bureau of Economic Research .
- Tavares, P., & Canuto, O. (17 de Agosto de 2018). *Banco Mundial Blogs*.
- Valverdi, N. M. (2017). Discriminación salarial: brecha salarial entre hombres y mujeres del mercado laboral paraguayo. *Universidad Nacional de Asunción*, 14.
- Váquiro, N. E. (2016). Brechas salariales por género en individuos con educación superior en Colombia: un análisis multi-metódico. *Universidad del Valle*, 56.
- World Economic Forum. (2018). *The global gender gap report*. Cologne: World Economic Forum.
- Yamada, G., Lavado, P., & Velarde, L. (2013). Habilidades no cognitivas y brecha de género salarial en el Perú. *Banco Central de Reserva del Perú*, 35.