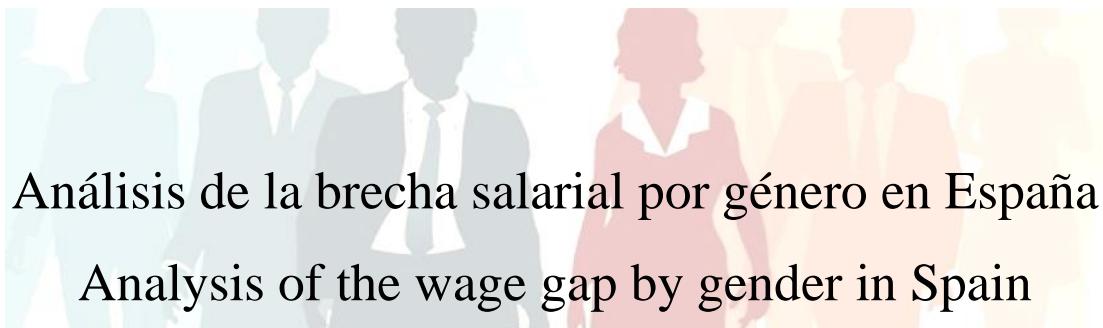




## Trabajo Fin de Grado



Análisis de la brecha salarial por género en España

Analysis of the wage gap by gender in Spain

Autor/es

Andrea González Cordón

Directores

**Monia Ben Kaabia**

**Jesús Clemente López**

Facultad de Economía y Empresa

2019

# ÍNDICE

INFORMACIÓN Y RESUMEN.....	- 1 -
1. INTRODUCCIÓN .....	- 2 -
2. REVISIÓN DE LA LITERATURA .....	- 4 -
3. ANÁLISIS DESCRIPTIVO DE LOS DATOS.....	- 9 -
3.1. Datos, variables y fuentes de información.....	- 9 -
3.1.1. Variables relacionadas con el trabajador:.....	- 10 -
3.1.2. Variables relacionadas con el mercado: .....	- 11 -
3.2. Análisis descriptivo .....	- 13 -
3.3. Análisis gráfico .....	- 15 -
4. APLICACIÓN ECONOMÉTRICA .....	- 17 -
4.1. Introducción .....	- 17 -
4.2 Método de la Variable ficticia .....	- 18 -
4.2.1. Especificación del modelo.....	- 18 -
4.2.2. Elección y definición de las variables .....	- 20 -
4.2.3. Estimación MCO y resultados.....	- 21 -
4.3. Descomposición de Oaxaca-Blinder .....	- 28 -
4.3.1. Metodología.....	- 29 -
4.3.2. Resultados.....	- 32 -
5. CONCLUSIONES .....	- 41 -
6. BIBLIOGRAFÍA .....	- 42 -
ANEXO I.....	- 43 -
ANEXO II .....	- 49 -
ANEXO III .....	- 60 -
ANEXO IV .....	- 68 -

## **INFORMACIÓN Y RESUMEN**

### **Resumen**

En España, al igual que en el resto del mundo, es un hecho que las mujeres cobran un salario menor al de los hombres. El presente trabajo, a partir de la Encuesta de Estructura Salarial de 2014, muestra que esta diferencia en el salario es de un 15-17% y se debe por una parte a las diferencias en las características de los individuos y por otra a la discriminación salarial por razones de género. Este resultado ha sido obtenido a través de estimaciones MCO utilizando la metodología de descomposición de Oaxaca (1973) y Blinder (1973).

Palabras clave: Discriminación salarial por género, brecha salarial, España.

### **Abstract**

In Spain, as in the rest of the world, it is a fact that women earn less than men. This study, based on the 2014 Wage Structure Survey, shows that this difference in salary is 15-17% and is due, on the one hand, to differences in the characteristics of people and, on the other, to wage discrimination based on gender. This result has been obtained through MCO estimates using the decomposition methodology of Oaxaca (1973) and Blinder (1973).

Key words: Wage discrimination by gender, wage gap, Spain

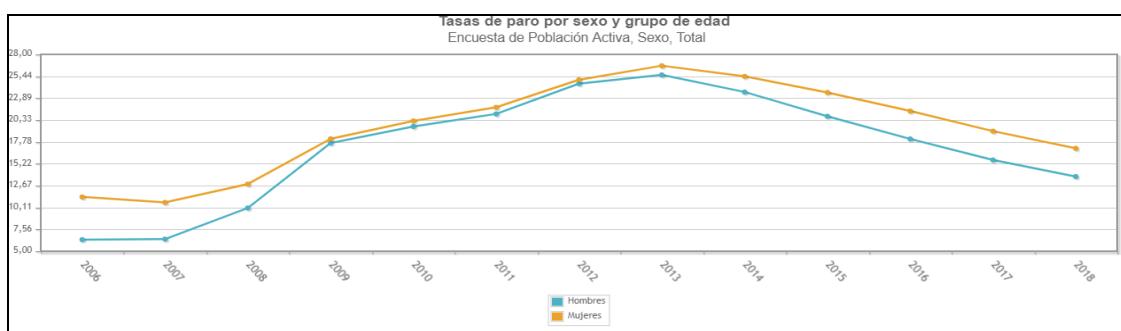
# 1. INTRODUCCIÓN

La discriminación salarial en contra de la mujer es un hecho que resuena en la actualidad de España a través de términos como “techo de cristal”<sup>1</sup> o de movimientos como el feminismo<sup>2</sup>.

La incorporación de la mujer al trabajo supuso un radical cambio en la vida de las mujeres puesto que empezaron a adquirir habilidades y a especializarse en actividades diferentes al cuidado de la familia y de la casa. Este proceso trajo consigo otros importantes cambios en la sociedad como es el descenso de la natalidad, el cambio de las estructuras familiares o el cambio de los roles dentro de las familias.

El gráfico 1.1 muestra la evolución de la tasa de desempleo de hombres y mujeres. Durante todo el periodo, la tasa de paro de las mujeres ha sido mayor que la de los hombres. Esta diferencia se redujo en la época de la crisis, a la vez que la tasa aumentaba para ambos géneros (25,6% para hombres y 26,67% para mujeres en 2013), especialmente en los hombres ya que el sector de la construcción, que fue uno de los más afectados, está mayoritariamente ocupado por hombres. Tras la crisis, las diferencias en la tasa de paro volvieron a aumentar en perjuicio de las mujeres.

**Gráfico 1.1. Evolución de la tasa de desempleo por género**



Fuente: INE

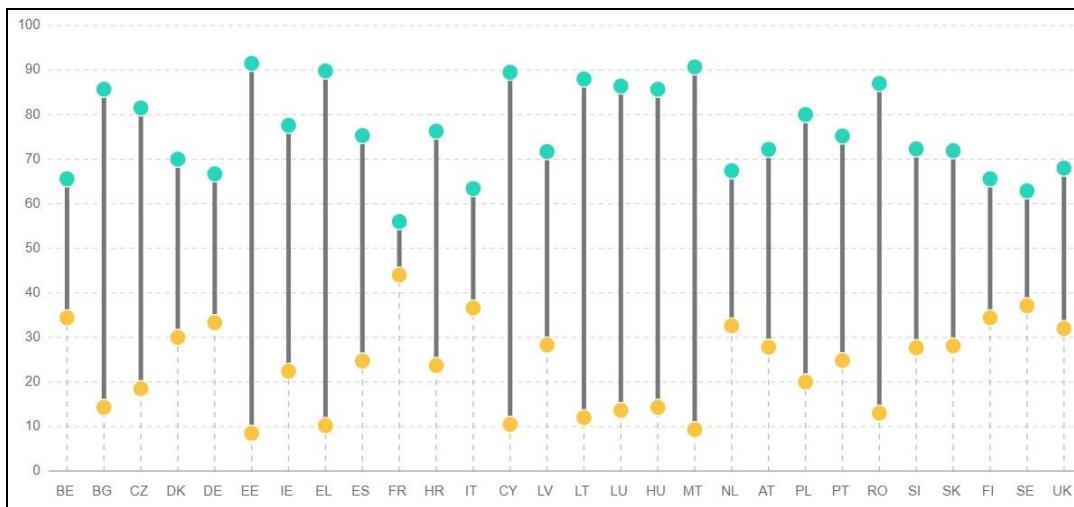
Este trabajo pretende poner sobre la mesa la realidad que sufren las mujeres que ven subestimadas sus capacidades y esfuerzo en sus puestos de trabajo y que merecen ser recompensadas de la misma manera que un hombre que hace el mismo trabajo. Además, es necesario romper ese “techo de cristal” que imposibilita la realización personal de las mujeres. El grafico 1.2 muestra el porcentaje de mujeres presentes en los

<sup>1</sup> Hace referencia a la limitación del ascenso laboral de las mujeres en el interior de las organizaciones.

<sup>2</sup> Movimiento que defiende la igualdad entre mujeres y hombres en todos los ámbitos.

consejos de administración de las mayores empresas que cotizan en bolsa. A pesar de que estas cifras han ido evolucionando a favor de la paridad de género, en todos los países, el número de hombres es mayor que el de mujeres.

**Gráfico 1.2. Distribución de los consejos de administración por género de las mayores empresas que cotizan en bolsa.**



Fuente: European Institute for Gender Equality

Nota: El punto azul corresponde al porcentaje de HOMBRES y el amarillo a MUJERES.

La discriminación ha tenido diversas teorías que intentan explicarla, como la teoría del capital humano, la teoría de la segmentación o teorías basadas en los gustos o en la información perfecta en las que más tarde nos adentraremos. Además es un tema objeto de estudio de numerosos artículos o publicaciones de carácter cuantitativo como los realizados por *Gardín y del Río (2009)*; *De la Rica (2017)* o los Estudios que lleva a cabo el Ministerio de Sanidad, Servicios Sociales e Igualdad, entre otros. Como veremos más tarde, todos ellos se inclinan hacia la idea de que existe discriminación salarial en contra de las mujeres.

Los objetivos principales de este trabajo son determinar el porcentaje de salario mensual por hora que cobran de menos las mujeres y ver como se descomponen los salarios para ver que parte de la brecha salarial se debe a las características personales de los individuos y que parte se debe a la discriminación. Además también es interesante ver en qué grupos de individuos, clasificados según las variables de análisis, se produce una mayor discriminación hacia las mujeres.

Para alcanzar estos objetivos se utiliza la Encuesta de Estructura Salarial (ESS) del año 2011, último año de publicación, que recoge datos sobre características de los

individuos tales como tipo de contrato, ocupación, edad o experiencia, además del salario. A través de modelos econométricos y de la descomposición de Oaxaca-Blinder obtendremos los resultados.

El trabajo está dividido en cuatro partes. Primeramente se realiza una revisión de la literatura con el objetivo de conocer las diferentes teorías que pretenden justificar la existencia de discriminación salarial en contra de las mujeres. A continuación se realiza un análisis descriptivo de los datos utilizados procedentes de la Encuesta de Estructura Salarial 2014. En el tercer apartado, se da a los datos una aplicación econométrica obteniendo unos modelos y la descomposición de Oaxaca-Blinder que darán los primeros resultados. Finalmente, a partir de los resultados obtenidos, se presentan las conclusiones.

## 2. REVISIÓN DE LA LITERATURA

Hay muchos estudios recientes sobre la brecha salarial por género en España, que a continuación analizaremos, y todos llegan a una misma conclusión: personas con las mismas características y el mismo trabajo perciben un salario menor cuando se es mujer que cuando se es hombre. En este apartado, en primer lugar analizaremos las diversas teorías que pretenden explicar la diferencia de salario por género y a continuación una serie de artículos/estudios relacionados con el tema.

Primeramente, la **teoría del capital humano** de *Becker* argumenta que la remuneración del trabajo viene determinada únicamente por la productividad marginal, independientemente del sexo. Según esto, dos personas con las mismas características, es decir, misma productividad, deberían cobrar un salario igual. Pero la realidad muestra que hombres y mujeres, que son igual de productivos, cobran un salario diferente, siendo mayor el de los hombres que el de las mujeres. Esta teoría ha sido puesta en duda por los numerosos estudios de *Blinder* (1973) y *Oaxaca* (1973), concluyendo que existe una parte de la diferencia salarial entre hombre y mujeres que no es explicada por las características de los mismos.

Por otro lado, la **teoría de la segmentación** del mercado de trabajo intenta dar explicación a lo que la teoría del capital humano no da. Para ello defiende que existen diferentes segmentos en el mercado y que en cada uno de ellos hay un sistema de

asignación de salarios diferente. Debido a la desigual distribución de hombre y mujeres en el mercado de trabajo, existe desigual distribución de salarios entre ellos.

Desde el lado de la Demanda, hay varias teorías que explican la discriminación. En primer lugar, tenemos **teorías basadas en los gustos**, propuestas por *Becker (1957)* y *Arrow (1973)*, en las que la discriminación se concibe como el gusto por el que el discriminador está dispuesto a pagar. Adoptan tres enfoques:

- 1- Quien discrimina es el empresario por tener prejuicios hacia el colectivo discriminado, las mujeres.
- 2- Quien discrimina son los trabajadores que soportan unos costes no pecuniarios por trabajar con mujeres.
- 3- Quien discrimina son los consumidores que soportan unos costes no pecuniarios por adquirir bienes o servicios producidos por una mujer.

En segundo lugar, nos encontramos con las **teorías basadas en la información imperfecta** de *Phelps et al (1972)*, en las que se toman como referencia característica promedio de los individuos de los distintos grupos y se asume que las características promedio de las mujeres son menores que las de los hombres. Por otro lado, tenemos el **modelo Overcrowding** en el que los distintos grupos de mujeres tienden a agruparse en determinados puestos de trabajo originando aglomeraciones que influyen en los salarios. Por último, nos encontramos con **nuevas perspectivas** tales como las actitudes frente al riesgo, hacia la competición o hacia la negociación. Estas perspectivas vienen a decir que las mujeres son más propensas a trabajos menos expuestos al riesgo, menos competitivos o en los que tengan que negociar para otras personas en lugar de para sí mismas, lo que conlleva que cobren un salario inferior.

El estudio de *Gardín y del Río (2009)* sobre los cambios experimentados en las diferencias salariales entre hombres y mujeres entre 1995 y 2002 concluye que, en este periodo, las diferencias se han reducido debido al aumento del número de mujeres con estudios universitarios. Sin embargo la brecha salarial aumenta en las mujeres que no tienen estudios universitarios. Se utiliza la metodología propuesta por *Jenkins (1994)* y desarrollada posteriormente por *Del Río, Gardín y Cantó (2006)*, permitiendo cuantificar la discriminación salarial total a partir de estimaciones individualizadas de la misma. Utiliza datos de la Encuesta de Estructura Salarial de 1995 y 2002.

Según el artículo de *Murillo y Simón (2014)* sobre diferencia salarial por género en España entre 2002 y 2010, la Gran Recesión ha aumentado la brecha salarial por género y ha propiciado el escenario, que hasta entonces no se venía dando, del “techo de cristal” debido al cual, las mujeres no pueden avanzar profesionalmente y ocupar cargos de alta responsabilidad. Dichos autores indican que, debido a la crisis, los hombres fueron los más perjudicados, lo que es coherente dado que uno de los sectores más afectados fue el de la construcción, ocupado mayoritariamente por hombres. Sin embargo, el empeoramiento de las condiciones laborales afectó más a las mujeres por ser un grupo más vulnerable. El artículo mencionado anteriormente utiliza dos métodos econométricos: una extensión de la metodología propuesta por *Juhn et al. (1991 y 1993)* adaptada a microdatos y que permite descomponer el diferencial del salario promedio entre hombres y mujeres y la metodología elaborada por *Fortin et al. (2011)* que proporciona una descomposición de las diferencias salariales entre hombres y mujeres a lo largo de la distribución salarial. Los datos que utiliza son los de la Encuesta de Estructura Salarial de 2002, 2006 y 2010.

Cabe destacar también el estudio realizado por *Genoveva, Santos y Pérez (2015)*, sobre la evolución y los factores socioeconómicos determinantes del empleo femenino en España. Las autores hablan de la existencia del “techo de cristal” ya mencionado anteriormente. Además, a pesar de que la Comisión Europea intentó combatir la brecha salarial a través de la "Estrategia para la igualdad entre mujeres y hombres 2010-2015", las diferencias siguen siendo significativas. Concluyen que el empleo femenino se recupera en 2014 y que las características óptimas para tener empleo son: ser joven, con un nivel de estudios alto, no tener hijos y trabajar a media jornada en el sector terciario. También ponen de manifiesto que, en general, las mujeres ocupan unos puestos de trabajo de menor cualificación que los que ocupan los hombres, que la mayoría de los empleos a tiempo parcial son cubiertos por mujeres y que la tasa de empleo femenina empieza a caer cuando las mujeres comienzan a tener hijos. La metodología utilizada se basa en la estimación de modelos Logit tomando como datos la Encuesta de Población Activa del INE con un tamaño muestral de 500 mujeres españolas.

El artículo *Cerviño (2018)* de sobre la desigualdad salarial de género en España en el contexto de la crisis económica y la recuperación, analiza la evolución de la brecha salarial en los años recientes y sus resultados muestran que la tendencia a la

disminución de la brecha salarial se vio afectada por la crisis, sobre todo a partir de 2010 pero que rápidamente reaccionó en 2016, situándose en un nivel inferior que en el que se situaba antes de la crisis. Con este artículo se puede corroborar que la brecha salarial sufre cambios de tendencia por lo que es necesario ahondar más en sus causas para poder combatir mejor el problema que representa. .

Es interesante también, el estudio llevado a cabo por *De la Rica (2017)* sobre la brecha de género en el mercado laboral español y su evolución a lo largo del ciclo de vida. En él, aparte de evidenciar la existencia de brecha salarial por género, se afirma que existe tanto en participación salarial como en salarios-hora. Además la brecha crece conforme aumenta la edad debido a que hay menos mujeres con estudios superiores y porque las mujeres reciben menos en relación a complementos variables debido a la existencia de segregación de género en determinados sectores.

Por otro lado, el análisis que lleva a cabo el Ministerio de Sanidad, Servicios Sociales e Igualdad (*De Lucio, Del Valle y Valero 2019*) sobre los determinantes de la brecha salarial por género en España utiliza los datos de la Base de Datos de Estructura Salarial y Competitividad, que combina información de la Encuesta de Estructura Salarial 2006 y de la base SABI (Sistema de Análisis de Balances Ibéricos). Este análisis se centra, no solo en la relación de los salarios con las características de los trabajadores, sino también en la relación de los mismos con las características de las empresas y el entorno en el que operan, es decir, la competitividad, ya que estos dos aspectos son relevantes en las políticas de igualdad retributiva. Utiliza técnicas econométricas para modelizar la variable de estudio. En un primer momento, a través del análisis descriptivo se obtiene que las mujeres cobran menos, independientemente de las características de la persona, las condiciones laborales o del entorno geográfico o empresarial. Además, afirma que la brecha salarial es mayor si se mide por rendimiento, siendo un 41,3%, que si se mide por ganancia media por hora, un 19,3%. Como en todos los estudios anteriores se llega a la conclusión de que, a través de los análisis econométricos, las mujeres cobran un 21% menos que los hombres simplemente por su condición sexual, por ser mujeres. Aplicando el método de Oaxaca-Blinder, que es el que utilizaremos en este trabajo, se concluye que de ese 21%, un 53,4% se debe únicamente al sexo y el resto se debe a factores no observables y no incluidos en el modelo. Es decir, realmente las mujeres cobran un 11% menos que los hombres en igualdad de condiciones. El estudio propone incidir más a través de las políticas, en los

sectores con convenio colectivo de ámbito más reducido, con mayor tamaño, en sociedades anónimas, en el sector de la construcción y con mayor grado de competitividad ya que son los segmentos con mayor discriminación salarial.

Por último, el artículo “Brechas Salariales de Género en España” de *Anghel, Conde-Ruiz y Marra (2019)*; utiliza datos de la Encuesta de Estructura Salarial (EES) española y europea. El artículo afirma que cuantas más mujeres haya en un país trabajando, más crecerá el país y que se produce el ya comentado efecto del “techo de cristal” dado que la brecha aumenta en la parte superior de la distribución salarial. La brecha salarial también **aumenta** con:

- La edad: los hombres entre 30-39 años, ganan un 11% más que las mujeres con dicha edad, y los de entre 50-59 años ganan un 15% más,
- La antigüedad en la empresa: los hombres con más de 7 años de antigüedad en la empresa ganan un 16,4% más que las mujeres con la misma antigüedad y características,
- Durabilidad del contrato: los hombres con contrato indefinido ganan un 14% más que las mujeres con el mismo contrato y los hombres con un contrato temporal ganan un 8% más,
- Tamaño de la empresa: los hombres ganan un 14,3% más que las mujeres en las empresas de más de 200 trabajadores y un 9,4% más en las empresas de menos de 10 trabajadores.

Y se **reduce** con:

- El nivel educativo: hombres con educación secundaria ganan un 14% más y los hombres universitarios un 12% más que las mujeres.

El artículo concluye diciendo que las mujeres son peor tratadas que los hombres en el mercado laboral y que eso puede suponer una pérdida de talento y crecimiento potencial para la economía del país.

Tras haber recopilado toda la información que estos estudios proporcionan, parece evidente que la brecha salarial por género sigue siendo un problema sin resolver. El artículo “Transparencia y brecha salarial” de *Camps (2018)*, cuestiona la universalidad del derecho al trabajo dada la existencia de la brecha salarial. En él se

propone como método para detectar estructuras empresariales discriminantes, a la transparencia.

La implementación de políticas favorables a la extinción de la brecha salarial por género debería ser un asunto de alta importancia en los programas electorales y, en general, en la sociedad, dado que el único objetivo que persiguen es la igualdad entre mujeres y hombres.

### **3. ANÁLISIS DESCRIPTIVO DE LOS DATOS**

#### **3.1. Datos, variables y fuentes de información**

Los datos empleados para esta investigación han sido obtenidos de La Encuesta de Estructura Salarial (ESS). Dicha fuente de información es una encuesta elaborada por el Instituto Nacional de Estadística (INE), pactada dentro de la Comunidad Europea y diseñada para obtener resultados acerca de la distribución de los salarios. El diseño de la ESS corresponde con una muestra en la cual se distinguen dos etapas de asalariados. En la primera etapa, se localizan las cuentas de cotización de la Seguridad Social, es decir, las empresas que tenemos en cuenta a la hora de realizar la encuesta; y en la segunda etapa se encuentra las cuentas de los trabajadores.

La Encuesta de Estructura Salarial (EES) es un método de obtención de información de obligado cumplimiento en todos los estados miembros de la Unión Europea con el fin de realizar comparaciones entre países. Se realiza cuatrienalmente y recoge información tanto de la ganancia media mensual por trabajador y su distribución por percentiles como de la ganancia media por hora trabajada y la composición del salario.

Este trabajo utiliza la EES de 2014 que toma una muestra de 209.436 individuos. Con el fin de analizar a los trabajadores, se utiliza la variable salario percibido por hora. Esta variable se ha obtenido a partir de los microdatos proporcionados por la encuesta. Las variables que aparecen recogidas en la base de datos se pueden dividir en dos grupos: las relacionadas con el trabajador y las relacionadas con el mercado.

### **3.1.1. Variables relacionadas con el trabajador:**

Dentro de este grupo distinguimos dos subdivisiones:

**1) Variables referentes al Capital Humano del trabajador:** tienen como finalidad analizar las capacidades de los trabajadores adquiridas con la educación, experiencia y preparación.

- Por un lado está el **nivel de estudios**, donde se distinguen las siete categorías correspondientes a la Clasificación Nacional de Educación 2014:

- Menos que primaria
- Educación primaria
- Primera etapa de educación secundaria
- Segunda etapa de educación secundaria
- Enseñanzas de formación profesional de grado superior y similares
- Diplomados universitarios y similares
- Licenciados y similares, y doctores universitarios

Esta variable recibe el nombre de “ESTU” y es una variable cualitativa que toma valores del 1 al 7 respectivamente.

- Por otro lado, también incluye información acerca de la **experiencia** del trabajador.

Esta variable es importante ya que se entiende que, durante ese periodo, el trabajador incorpora educación específica. Esta variable recibe el nombre de “ANOANTI” y es una variable cuantitativa medida en años.

**2) Variables referentes a las características propias del trabajador:** tienen como finalidad distinguir a los individuos unos de otros.

- Se diferencia por **sexo**. Esta variable cualitativa recibe el nombre de “SEXO” y toma el valor 1 si es Hombre y 6 Mujer. Es una variable importante para el objeto del estudio que trata este trabajo.

- Se diferencia por **edad** siguiendo los siguientes tramos:

- Menores o iguales a 19 años
- 20 – 29 años
- 30 – 39 años
- 40 – 49 años
- 50 – 59 años
- Mayores o iguales a 59 años

Esta variable cualitativa recibe el nombre de “AÑOS2” y toma valores del 1 al 6 respectivamente

- Se diferencia por **nacionalidad**. Esta variable cualitativa recibe el nombre de “TIPOPAIS” y toma valor 1 si es España y valor 0 si es Resto del Mundo.

-Se diferencia por la **unidad territorial** a la que pertenece el individuo para situarlo geográficamente:

- Noroeste: Galicia, Principado de Asturias y Cantabria.
- Noreste: País Vasco, Comunidad Foral de Navarra, La Rioja y Aragón.
- Com. Madrid: Comunidad de Madrid.
- Centro: Castilla y León, Castilla La Mancha y Extremadura.
- Este: Cataluña, Comunidad Valenciana e Islas Baleares.
- Sur: Andalucía, Murcia, Ceuta y Melilla
- Canarias: Islas Canarias.

Esta variable cualitativa recibe el nombre de “NUTS1” y toma valores del 1 al 7 respectivamente.

- Por último, se diferencia si el trabajador tiene algún tipo de **responsabilidad** o no dentro de la empresa. Es una variable cualitativa que recibe el nombre de “RESPONSA” y toma el valor 0 si no se tiene responsabilidad y el valor 1 si se tiene.

El objetivo de estas variables es determinar las características propias que presentan las empresas donde trabajan los individuos.

### 3.1.2. Variables relacionadas con el mercado:

1) **Variables relacionadas con el puesto de trabajo:** tiene como objetivo analizar las características de los empleos que ocupan los individuos.

- Se diferencia el **tipo de jornada**. Esta variable cualitativa recibe el nombre de “TIPOJOR” y toma el valor 1 si es a tiempo completo (8 horas diarias) y 2 si es a tiempo parcial (menos de 8 horas diarias).

- Se diferencia la **duración del contrato**. Esta variable cualitativa recibe el nombre de “TIPOCON” y toma el valor 1 si es indefinida y 2 si es determinada.

- Se diferencia por **Ocupación**, utilizando la Clasificación Nacional de Ocupaciones de 2011. Esta variable se divide en 17 categorías, correspondientes a los grupos principales del CNO-11. Podemos agrupar los distintos grupos de ocupaciones en cuatro categorías según su naturaleza y tipo de tareas que llevan a cabo. (Tabla 3.1, ver ANEXO I). Esta variable cualitativa recibe el nombre de “CNO1” y toma valores del 1 al 17 respectivamente.

- Se diferencia según la **Actividad Económica** en la que se encuentra el puesto de trabajo del individuo. Se distinguen 27 categorías según la CNACE. (Tabla 3.2, ver ANEXO I)

Esta variable cualitativa recibe el nombre de “CNACE” y toma valores del 1 al 27 respectivamente.

- Se diferencia el tipo de **regulación** de las relaciones laborales:

- Estatal sectorial
- Sectorial de ámbito inferior (autonómico, provincial, comarcal...)
- De empresa o grupo de empresas
- De centro de trabajo
- Otra forma de regulación.

Esta variable cualitativa recibe el nombre de “REGULACION” y toma valores del 1 al 5 respectivamente.

2) **Variables relacionadas con la empresa:** tiene como objetivo analizar las empresas en las que trabajan los individuos:

- Se diferencia el **tipo de mercado** donde la empresa destina su producción:

- Local o Regional
- Nacional
- Unión Europea
- Mundial

Esta arable cualitativa recibe el nombre de “MERCADO” y toma valores del 1 al 4 respectivamente

- Se diferencia el **tipo de empresa**. Esta variable cualitativa recibe el nombre de “CONTROL” y toma el valor 1 si es pública y 2 si es privada.

- Se diferencia el **tamaño de la empresa**, agrupando según el número de trabajadores:

- Incluye todos los estratos
- De 1 a 49 trabajadores
- De 50 a 199 trabajadores
- Más de 200 trabajadores
- Incluye estrato 2 y 3

Esta variable recibe el nombre de “ESTRATO2” y toma valores del 0 al 4 respectivamente.

### **3.2. Análisis descriptivo**

Del total de la muestra, el 57,27% son Hombres mientras que el 42,73% son Mujeres. El 95% de los individuos son españoles y viven repartidos por todo el territorio español, destacando el Este donde se ubica el 27,2% de la muestra. Son individuos mayoritariamente adultos ya que el 32,1% está entre los 30 y 39 años y el 31,1% entre los 40 y 49 años y con un nivel de estudios secundarios (46,9% de los hombres y 44,9% de las mujeres).

En relación a la actividad económica que desempeñan, la distribución es bastante igualitaria destacando para las mujeres las actividades sanitarias y de servicios sociales con un 14,5% y para los hombres la construcción con un 9,3%. En cuanto a la ocupación, las mujeres son mayoritariamente técnicas o profesionales de apoyo (16,9%) y empleadas de oficina que no atienden al público (12,7%). Los hombres son también mayoritariamente técnicos o profesionales de apoyo (18,6%) y trabajadores cualificados en industrias manufactureras (14,1%).

Casi la mitad de los individuos trabajan en empresas de 200 o más trabajadores (41,6%), privadas (84,3%) y con regulación autonómica, provincial o comarcal (35,8%). Las mujeres destacan más en empresas locales o regionales (43,1%) mientras que los hombres en empresas nacionales (41,1%). Tanto unos como otras no tienen responsabilidad dentro de la empresa. El 82,4% de los individuos trabajan a jornada completa (el 90,2% de los hombres y el 71,9% de las mujeres) y el 79,5% tienen un contrato indefinido (el 80,1% de los hombres y el 78,7% de las mujeres). Por último, en relación a la experiencia, el 39,97% de la muestra tiene menos de 6 años de experiencia laboral.

La tabla 3.3 del ANEXO I, recoge la distribución del salario por hora de la muestra según las características de interés.

Se observa que el salario por hora aumenta tanto para hombres como para mujeres a medida que el nivel de estudios es superior. Sin embargo, la brecha salarial y el nivel de estudios tienen una relación positiva, a medida que aumenta el nivel de estudios, la brecha salarial en valores absolutos también aumenta (las mujeres que no tienen ningún tipo de estudios cobran 1,82 euros por hora menos que los hombres; y las mujeres licenciadas o doctoradas cobran 5,26 euros por hora menos que los hombres). Vemos que con la edad ocurre lo mismo, a medida que aumenta la edad, hombres y mujeres cobran más pero la brecha salarial aumenta.

También se observa que tanto hombres como mujeres cobran más por hora en España que en el Resto del Mundo pero la brecha salarial es mayor en España (2,53) que en el Resto del Mundo (1,93). Dentro del territorio español, los salarios por hora están distribuidos de manera igualitaria entre todas las zonas entre 9 y 14 euros la hora tanto para mujeres como hombres, siendo el salario de los hombre siempre mayor que el de las mujeres. La mayor brecha se da en la Comunidad de Madrid (3,66 euros por hora de diferencia). Los trabajadores y trabajadoras que trabajan jornada completa cobran más que los que tiene jornada parcial, sin embargo, la brecha salarial es 0,4 euros por hora mayor en la jornada parcial. Los individuos que tienen contrato indefinido tienen un salario por hora mayor que los que tienen contrato determinado, pero en los primeros la brecha salarial es 2,07 euros mayor.

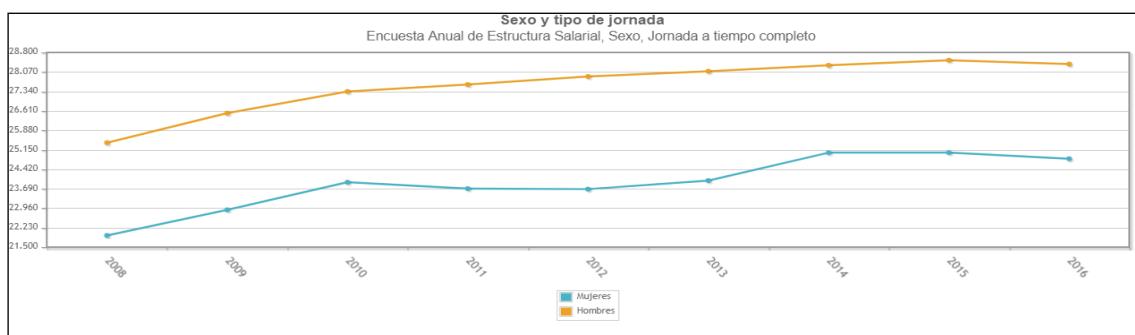
En relación a la actividad económica, la brecha siempre es positiva (los hombres cobran más que las mujeres) excepto en Suministro de energía eléctrica, gas, vapor y aire acondicionado donde las mujeres cobran 0,11 euros por hora más y en Actividades administrativas y servicios auxiliares donde las mujeres cobran 0,92 euros por hora más. Cantidades relativamente inferiores a las de las demás actividades económicas en las que los hombres cobran entre 1 y 4 euros más, destacando la Industria manufacturera de artes gráficas y reproducción de soportes grabados donde los hombres cobran 6,03 euros por hora más. En cuanto a las ocupaciones, la brecha salarial siempre es positiva a favor de los hombres. Destaca sobre todo los directores y gerentes que cobran casi 8 euros más.

Por último, en relación a las características de las empresas, vemos que los salarios y la brecha salarial aumentan a medida que el ámbito de la empresa es mayor. La brecha es de 1,26 euros por hora mayor en las empresas locales o regionales, sin embargo en las empresas multinacionales la brecha es de 3,74 euros por hora. Ocurre lo mismo a medida que la empresa aumenta de tamaño. Cuantos más trabajadores tiene la empresa, mayores son los salarios pero mayor es la brecha salarial entre hombres y mujeres. En cuanto al capital de la empresa, vemos que los salarios son mayores en las empresas públicas que en las privadas y además la brecha salarial es menor (1,87 euros por hora en las públicas frente a 2,84 euros por hora en las privadas).

### 3.3. Análisis gráfico

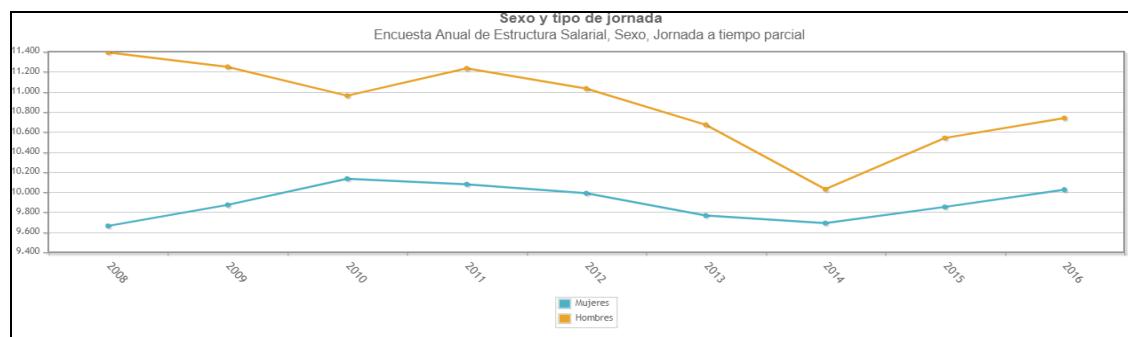
La brecha salarial entre hombres y mujeres se ha mantenido constante desde 2008 a 2016 para los trabajadores con jornada a tiempo completo (Gráfico 3.1). Sin embargo, para los trabajadores con jornada parcial, la brecha salarial disminuyó hasta 2010 lo que pudo ser causa de las consecuencias de la crisis sobre todo en el sector de la construcción, mayormente ocupado por hombres. Desde entonces, aumento un poco para después volver a reducirse hasta 2014. Desde entonces parece que se ha mantenido constante. (Gráfico 3.2)

**Gráfico 3.1. Brecha salarial por género entre trabajadores con jornada a tiempo completo.**



Fuente: elaboración propia a partir de la Encuesta Anual de Estructura Salarial (INE)

**Gráfico 3.2. Brecha salarial por género entre trabajadores con jornada a tiempo parcial.**

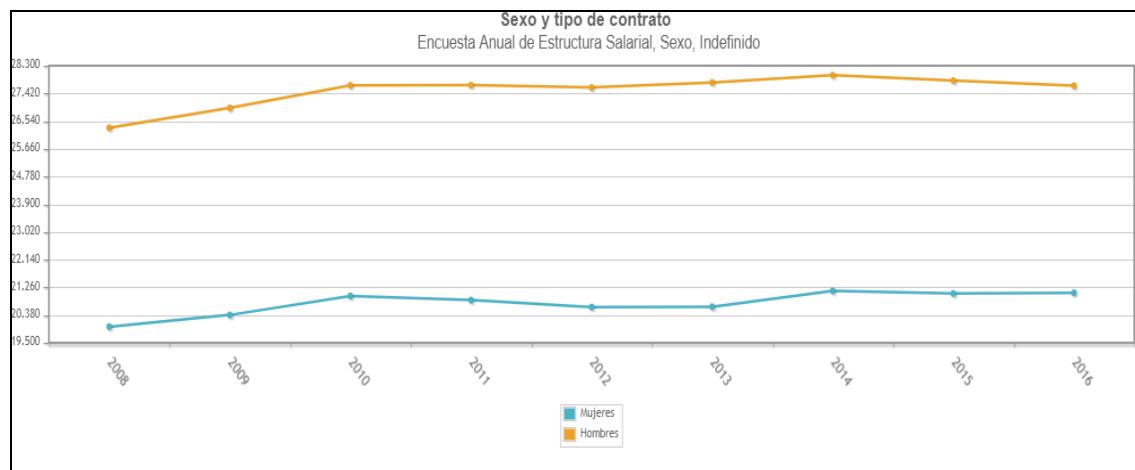


Fuente: elaboración propia a partir de la Encuesta Anual de Estructura Salarial (INE)

En relación al tipo de contrato, la brecha salarial entre hombres y mujeres no se ve afectada ya que se mantiene constante durante el periodo tanto para contratos indefinidos (Gráfico 3.3) como para temporales (Gráfico 3.4). Pero hay que destacar que mientras que en los contratos indefinidos, los salarios de hombres y mujeres no han sufrido grandes variaciones, en los contratos temporales, el salario tenía una tendencia

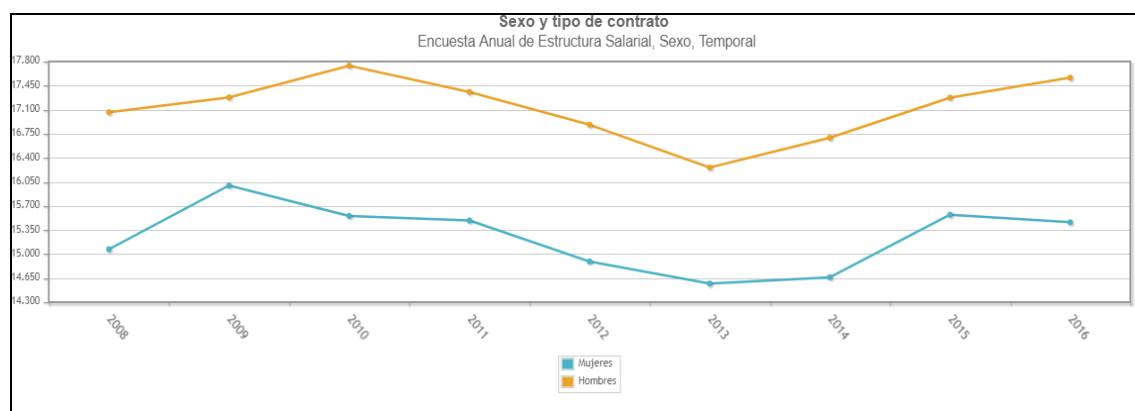
creciente hasta 2010 y a partir de entonces, debido a la crisis, los salarios caen para posteriormente recuperarse a partir de 2013.

**Gráfico 3.3. Brecha salarial por género entre trabajadores con contrato indefinido.**



Fuente: elaboración propia a partir de la Encuesta Anual de Estructura Salarial (INE)

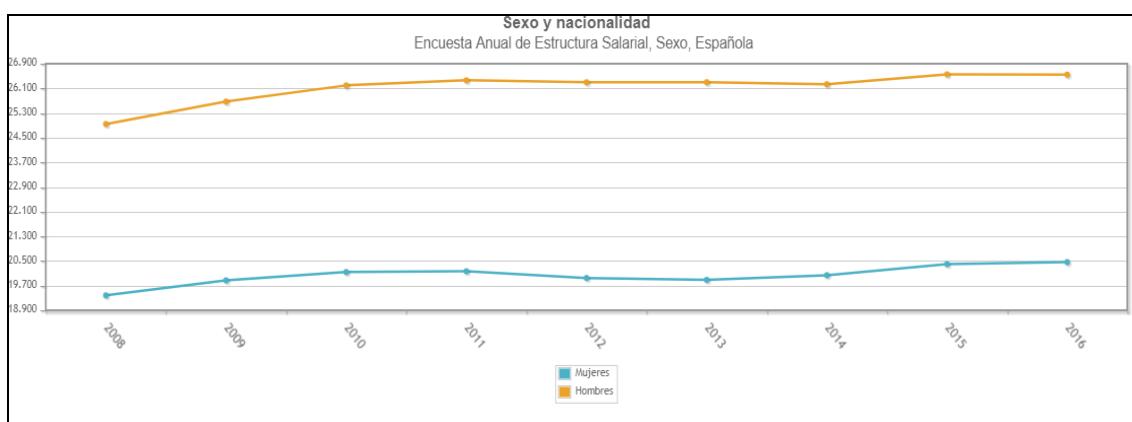
**Gráfico 3.4. Brecha salarial por género entre trabajadores con contrato temporal.**



Fuente: elaboración propia a partir de la Encuesta Anual de Estructura Salarial (INE)

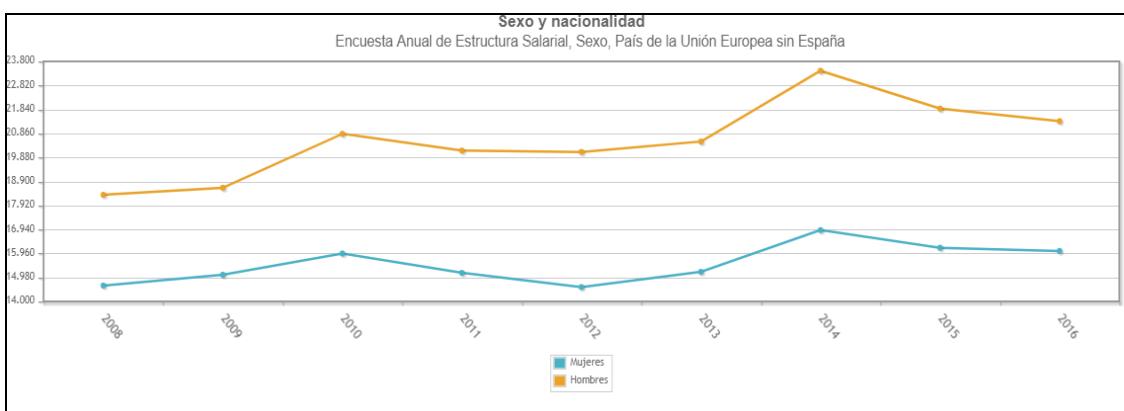
Por último, la brecha salarial se mantiene constante en España durante el periodo (Gráfico 3.5) al igual que en el resto de países de la Unión Europea (Gráfico 3.6). Hay que destacar que tanto en 2010 como el 2014, en el resto de los países de la Unión Europea hubo un repunte de los salarios de hombres y mujeres cosa que no sucedió en España, donde los salarios se han mantenido bastante constantes, existiendo una diferencia de aproximadamente 6.000 € entre hombres y mujeres. En el resto de países de la Unión Europea esta es diferencia es menor, aproximadamente 4.000 €, aunque continua siendo significante.

**Gráfico 3.5. Brecha salarial por género entre trabajadores españoles.**



Fuente: elaboración propia a partir de la Encuesta Anual de Estructura Salarial (INE)

**Gráfico 3.6. Brecha salarial por género entre trabajadores no españoles.**



Fuente: elaboración propia a partir de la Encuesta Anual de Estructura Salarial (INE)

## 4. APLICACIÓN ECONOMÉTRICA

### 4.1. Introducción

Como ya se ha comentado en la introducción, el objetivo principal del presente trabajo se centra en analizar la discriminación salarial entre hombres y mujeres en el mercado laboral español.

En la literatura económica el concepto de discriminación salarial hace referencia a la parte de brecha salarial que no viene explicada por diferencias en los puestos de trabajo en las características productiva entre dos grupos de trabajadores (A y B). Utilizando la información de la ESS correspondiente al año 2014 (última encuesta disponible), en el presente trabajo se van a utilizar dos métodos para tratar de explicar la brecha salarial entre hombres y mujeres. El primer método se basa en la estimación

Mínimos Cuadrados Ordinaria (MCO), mediante el uso de la variable ficticia de género. El segundo método se Oaxaca y Blinder.

## 4.2 Método de la Variable ficticia

### 4.2.1. Especificación del modelo

El método de la variable ficticia es uno de los métodos más sencillos para medir la discriminación salarial entre hombres y mujeres. En concreto, se trata de estimar un modelo econométrico que trata de explicar el salario percibido en función de un conjunto de variables explicativas, entre las cuales debe aparecer la variable ficticia que recoja el sexo del trabajador. Para ello, tenemos que obtener una correcta especificación del modelo econométrico que trata de explicar el salario.

*Mincer (1974)* fue quien primero analizó la relación entre la distribución de las retribuciones y el capital humano; es el responsable de desarrollar el análisis empírico de la relación entre capital humano y distribución personal de ingresos, así como del concepto de tasa de rentabilidad de la educación, que, sin duda, constituyó la piedra angular de un gran número de investigaciones en esta área (Barceinas et. al, 2001: 5).

La ecuación de regresión de Mincer, llamada “función de ingresos”, incluye el logaritmo del ingreso como variable dependiente y la educación y los años de experiencia como variables explicativas. Esta función ha proporcionado las bases para un amplio conjunto de investigación empírica en relación con el nivel y distribución del ciclo vital de ingresos y los rendimientos de la educación (Barceinas et. al, 2001:8).

El modelo econométrico más utilizado en las aplicaciones empíricas es el siguiente:

$$\ln(\text{Salario}_i) = \beta_0 + \beta_1 \text{Educ}_i + \beta_2 \text{EXP}_i + \beta_3 \text{EXP}_i^2 + u_i \quad (1)$$

Donde:

- **Salario** es el salario percibido. En el presente trabajo se ha utilizado el salario mensual por hora trabajada para analizar los trabajadores con diferentes jornadas en condiciones de igualdad. Dicho salario se ha calculado como la ganancia mensual dividida entre las horas trabajadas al mes de referencia.
- Educ son los años de educación,
- EXP es la experiencia laboral,
- EXP<sup>2</sup> es la experiencia al cuadrado

- $u_i$  es la perturbación aleatoria del modelo econométrico que debería cumplir los supuestos básicos para que el modelo sea especificado correctamente, es decir,  $\varepsilon_t \sim i.i.N(0, \sigma^2)$  como se conoce en econometría un ruido blanco.

La introducción del cuadrado de la experiencia como explicativa en el modelo trata de captar la no linealidad del perfil edad-ingresos.

Las  $\beta_i$  son los parámetros del modelo y que hay que estimar. En concreto,  $\beta_0$  es el término constante, y representa el logaritmo del salario de un individuo que no tiene educación ni experiencia. En la literatura clásica, el coeficiente  $\beta_1$  se ha denominado “la tasa de rendimiento privada de la educación”.

Teóricamente,  $\beta_1$  y  $\beta_2$  deben ser positivos, y  $\beta_3$  debe ser negativo. Es decir la relación salario experiencia tendría una forma de U invertida. Crece al principio y a partir de un determinado momento, el efecto de la experiencia vuelve a ser negativo (Barceinas, 2001 y Areiza et. al, 2004).

El modelo especificado en la ecuación (1), sólo capta la relación entre salario, educación y experiencia. Por tanto, si el objetivo final es tratar de analizar la brecha salarial entre hombre y mujeres, el modelo debe de incluir otras variables explicativas. Se supone que el salario que percibe un trabajador es igual a su productividad y depende de las características personales, del capital humano, del puesto de trabajo y de la empresa donde desarrolla su trabajo. En base a eso, el modelo queda especificado de la siguiente manera:

$$\ln(\text{Salario}_i) = \beta_0 + \beta_1 \text{Edu}_i + \beta_2 \text{Exp}_i + \beta_3 \text{Exp}_i^2 + X_i^\top \lambda + \delta \text{Genero}_i + u_i \quad (2)$$

Donde:

- $X_i$  es el conjunto de otras variables explicativas (como nivel de estudios, localidad geográfica, procedencia, etc.) que puedan influir en el salario aparte de la Edu, Exp y Género.
- Género es una variable cualitativa que toma el valor 1 si el trabajador  $i$  es de género masculino y 0 en caso contrario.

Según dicha especificación, el coeficiente  $\delta$  mide la diferencia en el salario entre un hombre y una mujer. Si su valor estimado es positivo y estadísticamente significativo, se dice que existe discriminación salarial en contra de la mujer. Es decir, para el mismo valor de los demás factores, los hombres ganan un  $100\hat{\delta}$  más que las mujeres.

#### 4.2.2. Elección y definición de las variables

Como ya hemos indicado la variable dependiente del modelo es el logaritmo neperiano del salario por hora. Las variables independientes para explicar el salario se han dividido en dos grupos: las relacionadas con el trabajador (Tabla 4.1) y las referentes al mercado laboral o empresa (Tabla 4.2).

**Tabla 4.1. Variables explicativas del modelo relacionadas con el trabajador.**

VARIABLES relacionadas con el trabajador
1) Experiencia: (ANOANTI ) Es una variable cuantitativa que define la antigüedad en el puesto de trabajo
2) (Experiencia) <sup>2</sup> : (ANOANTI) <sup>2</sup>
3) Edad: es una variable cualitativa con 6 categorías: ANOS1: toma el valor 1 si $EDAD \leq 19$ años ANOS2: toma el valor 1 si $20 \leq EDAD \leq 29$ años ANOS3: toma el valor 1 si $30 \leq EDAD \leq 39$ años ANOS4: toma el valor 1 si $40 \leq EDAD \leq 49$ años ANOS5: toma el valor 1 si $50 \leq EDAD < 59$ años ANOS6: toma el valor 1 si $EDAD \geq 59$ años
4) Localidad a la que pertenece el trabajador: Es una variable cualitativa se divide en 7 categorías: NUTS1_1: Noroeste: Galicia, Principado de Asturias y Cantabria. NUTS1_2: Noreste: País Vasco, Comunidad Foral de Navarra, La Rioja y Aragón. NUTS1_3: Comunidad de Madrid. NUTS1_4: Centro: Castilla y León, Castilla La Mancha y Extremadura. NUTS1_5: Este: Cataluña, Comunidad Valenciana e Islas Baleares. NUTS1_6: Sur: Andalucía, Murcia, Ceuta y Melilla NUTS1_7: Canarias: Islas Canarias
5) Estudios: Es una variable está dividida en 7 categorías: ESTU1: Menos que primaria ESTU2: Educación primaria ESTU3: Primera etapa de educación secundaria ESTU4: Segunda etapa de educación secundaria ESTU5: Enseñanzas de formación profesional de grado superior y similares ESTU6: Diplomados universitarios y similares ESTU7: Licenciados y similares, y doctores universitarios
6) Nacionalidad del trabajador: TIPOPAIS Es una variable de 2 categoría y que toma el valor 1 si es Español y 0 si es extranjero

Fuente: elaboración propia.

Dentro de las variables relacionadas con el puesto de trabajo hemos considerado el tipo de jornada, la duración del contrato, la ocupación, el tamaño de la empresa, empresa pública o privada y el tipo de responsabilidad en el puesto de trabajo. Estas

variables nos permiten analizar cómo influyen las características propias del puesto de trabajo que desempeña el empleado tanto sobre el salario percibido así como sobre la brecha salarial y la discriminación por género. En la siguiente tabla (Tabla 4.2) se detallan las variables utilizadas como explicativas dentro de esta categoría.

**Tabla 4.2. Variables explicativas del modelo relacionadas con el puesto de trabajo.**

VARIABLES relacionadas con el puesto de trabajo
1) Tipo de contrato: TIPOCON Es una variable de dos categorías: duración indefinida (valor 1) y duración temporal (valor 0)
2) Tipo de jornada: TIPOJOR Es una variable de dos categorías: tiempo completo (valor 1) y tiempo parcial (valor 0)
3) Ocupación: CNO1 Se detallan en la tabla 3.1 del ANEXO
4) Tamaño de la empresa: ESTRATO2. La variable de tamaño de la empresa se divide en 3 categorías ESTRATO21: De 1 a 49 trabajadores ESTRATO22: De 50 a 199 trabajadores ESTRATO23: Más de 200 trabajadores
5) Tipo de empresa: CONTROL Es una variable de dos categorías: toma el valor 1 si es pública y 0 si es privada
6) Tipo de mercado: MERCADO La variable se divide en 4 categorías según el <b>tipo de mercado</b> donde la empresa destina su producción: MERCADO1: Local o Regional MERCADO2: Nacional MERCADO3: Unión Europea MERCADO4: Mundial

Fuente: elaboración propia.

A la hora de estimar el modelo y con el objetivo de evitar la trampa de las variables ficticias, para cada una de las variables cualitativas, hay que elegir una de la categoría como variable de control. Hemos decidido en cada caso dejar la primera de las categorías como la categoría de referencia o control.

#### 4.2.3. Estimación MCO y resultados

El modelo especificado en la ecuación (2) se pudo estimar por Mínimos Cuadrados Ordinarios (MCO). Sin embargo, para que los estimadores cumplan las propiedades deseadas (insesgadez y consistencia), las perturbaciones del modelo deben ser ruido blanco. Por tanto, una vez estimado el modelo, es necesario verificar que los residuos del modelo se comportan como un ruido blanco. Al tratarse de un modelo con datos de corte transversal, las hipótesis que debemos verificar es la homoscedasticidad.

La no homoscedasticidad se conoce como heteroscedasticidad que es un fenómeno **frecuente en datos de sección cruzada.** Todos los estadísticos se basan en la contrastación de la hipótesis nula de ausencia de heteroscedasticidad, éstos se pueden clasificar en dos categorías en función de la información recogida en la hipótesis alternativa. En el primer grupo sobresalen aquellos que sugieren la forma funcional de la heteroscedasticidad cuando se rechaza la hipótesis nula. Dentro de este grupo se incluyen el contraste de Glesjer (1969) y el del Multiplicador de Lagrange (LM) desarrollado por Breusch y Pagan (1979, 1980). Dentro del segundo grupo se encuentra el contraste de Goldfeld y Quandt (1965) (GQ) y de White (1980) que es bastante general ya que no requiere ningún supuesto *ad hoc* sobre la forma funcional de la heteroscedasticidad. En este sentido, White sostiene que éste es un contraste general de especificación errónea del modelo ya que la hipótesis nula consiste en que el término de la perturbación es homoscedástica e independiente de los regresores, y que la especificación del modelo es la correcta.

En este trabajo el contraste de Breusch-Pagan será utilizado para verificar el cumplimiento de la hipótesis de homoscedasticidad. El no rechazar las hipótesis nulas implica que el modelo está correctamente especificado. Por el contrario, rechazar  $H_0$  se traduce en un síntoma de mala especificación del modelo sea por errores de especificación o bien por la propia naturaleza de los datos. Sin embargo, a pesar que las estimaciones de los parámetros bajo la presencia de heteroscedasticidad son consistentes, la matriz de varianzas y covarianzas de los estimadores es inconsistente. Este último, se traduce en que los contrastes de hipótesis no son válidos. Una posible solución a este problema sería el uso de la matriz de varianzas y covarianzas consistentes bajo heteroscedasticidad aplicando el método de estimaciones robustas propuesto por White (1980).

Los resultados de la estimación del modelo por MCO se recogen en la Tabla 4.3 (ver ANEXO III). Una vez estimado el modelo, a continuación, se ha contrastado si los residuos del modelo estimado se aproximan a un ruido blanco, es decir, si cumplen la hipótesis de homoscedasticidad y normalidad. Los resultados obtenidos de los contrastes de validación aparecen recogidos en la tabla 4.4. En el caso del contraste de Breusch-Pagan, en todos los casos se han considerado todas las explicativas del modelo como las posibles variables que pueden generar la heteroscedatidad

**Tabla 4.4. Resultados de los contrastes de heterocedasticidad y nomarilidad.**

CONTRASTE	HIPOTESIS NULA	ESTADÍSTICO	CONCLUSIÓN
Homocedasticidad: Breuch-Pagan	Ho: Homocedasticidad	LM=17559,7 P-valor=0	Rechazo H0
Normalidad	Ho: Normalidad	X <sup>2</sup> = 28450,2 P-valor=0	Rechazo H0

Fuente: elaboración propia.

Los valores obtenidos del estadístico de Breusch-Pagan indican que la hipótesis de homoscedasticidad se rechaza al nivel de significación del 5%. Por otra parte, los resultados obtenidos a partir del contraste de normalidad rechazan la hipótesis nula de normalidad. La heteroscedasticidad, generalmente, sea viene generada por una mala especificación del modelo, o por la propia calidad de los datos. Observando detenidamente los residuos del modelo estimado, se observa la presencia de un gran número de observaciones que pueden considerarse atípicos. En definitiva, todos los intentos en tratar de solucionar el problema no han tenido éxito.

Teniendo en cuenta que uno de nuestros objetivos es analizar la discriminación salarial y obtener más adelante la descomposición, y para proceder a la siguiente etapa de la metodología econométrica y con el propósito de que los distintos contrastes de hipótesis sobre los parámetros del modelo sean válidos<sup>3</sup> se ha decidido utilizar las desviaciones típicas robustas a la heteroscedasticidad siguiendo la metodología de White (1980).

En cuanto a la bondad de ajuste del modelo, se ha calculado el coeficiente de determinación. Los valores calculados aparecen recogidos en la tabla 4.5. Como se puede observar los valores de los coeficientes de determinación no son muy elevados, algo bastante común cuando se trabaja con datos de corte transversal (Wooldridge, 2001). A pesar de ello, se puede considerar que los valores obtenidos en este trabajo son bastantes aceptables ya que oscilan en torno a 0,50. Estos resultados son similares a los obtenidos en otros trabajos con datos de corte transversal (Gracia, 1994, Caballero y Uriel, 1989) y en algunos casos hasta resultan más elevados.

---

<sup>3</sup> Debemos tener en cuenta que en presencia de problemas de heteroscedasticidad, aunque los estimadores MCO de los parámetros siguen siendo insesgados, la inferencia estadística utilizando la matriz de varianzas y covarianzas MCO no son válidos (Para más detalles ver Wooldridge, 2001).

**Tabla 4.5. Resultados del coeficiente de determinación.**

SUMA TOTAL	SUMA EXPLICADA	SUMA RESIDUAL	R <sup>2</sup>
51152,9	25912,8	25240	<b>0,56</b>

Fuente: elaboración propia.

Tras comprobar la correcta especificación del modelo a partir de los contrastes de homoscedasticidad, la siguiente etapa consiste en llevar a cabo ciertos contrastes de significatividad tanto de tipo individual (t-ratios) como de tipo conjunta de las variables explicativas incluidas en los dos modelos.

Para los contrastes de significatividad individual de los parámetros se han utilizado los t-ratios robustos a la heteroscedasticidad. En concreto, se trata de contrastar para cada uno de los parámetros del modelo especificado la siguiente hipótesis nula y alternativa:

$$H_0: \beta_i = 0 \Leftrightarrow H_0: \beta_i \text{ individualmente no significativo}$$

$$H_a: \beta_i \neq 0 \Leftrightarrow H_a: \beta_i \text{ individualmente significativo}$$

El estadístico de contraste es el t-ratio calculado de la siguiente forma:

$$t\text{-ratio} = \frac{\hat{\beta}_i}{\hat{\sigma}_{\hat{\beta}_i}} \sim t\text{-tudent}$$

Siendo los siguientes resultados a partir de estos contrastes, los más destacados:

- 1) El coeficiente que acompaña a la variable SEXO resulta ser positivo e individualmente significativo al 5% de nivel de significación. Es decir, para el mismo valor de los demás factores, los hombres ganan un 15,72% más que las mujeres.
- 2) Los rendimientos de la variable experiencia son positivos y los de su cuadrado son negativos. Esto representa la concavidad de la función de ingresos respecto de la experiencia (U invertida). Los individuos, por cada año adicional de experiencia, manteniendo el resto de variables constantes, incrementarán su salario mensual por hora pero cuando lleguen a una determinada edad de experiencia, sus ingresos se reducen (79 años). Es un dato poco relevante ya que teniendo en cuenta que la edad mínima legal a partir de la cual se puede comenzar a trabajar son 16 años, una persona con 79 años de experiencia quiere decir que como mínimo tiene 95 años y ya no está en condiciones de trabajar. Por lo que aumentos en los años de experiencia conllevan aumentos en el salario mensual por hora.

- 3) Todos los coeficientes que acompañan a las variables ficticias (excepto la DESTUD\_2) del nivel de educación son individualmente y conjuntamente significativas al 5%. Teniendo en cuenta que la categoría de control es la de nivel de estudio menos que primaria, el hecho de que todos los parámetros son positivos significa que a mayor nivel de estudio, mayor es el salario mensual por hora. Por ejemplo si nos fijamos en la variable ficticia DESTU\_5 que representa el nivel de estudios de formación profesional superior, se ve que los individuos con este nivel de estudios cobran un 13% más que los individuos sin estudios. La diferencia en el salario entre los individuos con menor nivel de estudios y los de mayor nivel de estudios es un 30% favorable a los segundos.
- 4) Si nos fijamos en la edad, todos los coeficientes son individualmente y conjuntamente significativos (excepto DANOS2\_2). Todos los grupos de edades reciben un porcentaje mayor de salario mensual por hora respecto al grupo de control que son los menores de 19 años.
- 5) Tomando como variable de control el Noroeste (Galicia, Principado de Asturias y Cantabria) de España, los resultados muestran que en el Noreste, Comunidad de Madrid, Este, Sur de España y Canarias se cobra un porcentaje mayor que en el Noroeste (10,7%, 7%, 6,5%, 3,4% y 0,8% respectivamente). Sin embargo en el Centro se cobra un porcentaje menor (1%). Además los individuos del resto del mundo cobran un 1,2% más que los españoles.
- 6) Para acabar esta sección, los resultados relacionados con la responsabilidad que el individuo tiene en la empresa, muestran que se cobra un 15% menos si no se tiene responsabilidad que si se tiene, lo cual es coherente. A mayor responsabilidad, mayor salario.

#### Resultados relacionados con las características del mercado

- 7) En relación con el tipo de ocupación los resultados muestran que un individuo cobra un porcentaje menor de salario mensual por hora en cualquier ocupación respecto a la de control que es la de directores y gerentes. Destaca que la ocupación militar cobra un 115% menos.
- 8) Respecto al resto de variables que representan las características del mercado se observa que se cobra un porcentaje **mayor** de salario mensual por hora cuanto más grande es la empresa, es decir, cuantos más trabajadores tiene contratados (7% más las empresas que tienen de 50 a 199 trabajadores y 12% más las

empresas con más de 200 trabajadores, respecto a las que tienen entre 1 y 49 trabajadores); cuanto más grande es el ámbito a que la empresa destina su producción (3% más las nacionales, 9% más las de la Unión Europea y 13% más las mundiales, respecto a las locales o regionales). Se cobra un 5,5% menos de salario mensual por hora si se trabaja en una empresa privada que si se trabaja en una pública; un 4,5% menos si se tiene un contrato determinado que si se tiene indefinido, y un 5,2% menos si se trabaja a jornada a tiempo parcial que si se trabaja a jornada completa.

Una vez comentados los resultados de la estimación del modelo, a continuación realizamos el contraste de Chow para contrastar si la Recta de Regresión Muestral de los hombres es la misma que de las mujeres. La tabla 4.6 del ANEXO III recoge el modelo en el que se realiza el contraste de Chow para comprobar si existe o no diferencia estructural con respecto a la variable SEXO. Los resultados de dicho contraste rechazan la hipótesis nula ( $H_0$ : no hay diferencia estructural) con una  $F(46, 209344) = 211,541$  y un  $p$ -valor = 0,0.

A la vista de dicho resultados, se ha decidido estimar un modelo para el grupo de hombres y otro para el de las mujeres. Las tablas 4.7 y 4.8 del ANEXO III recogen los resultados obtenidos para hombres y mujeres respectivamente. A partir de dichas estimaciones, la discriminación salarial entre hombres y mujeres se puede analizar comparando los parámetros estimados para cada uno de los grupos (Hombres y mujeres). Dichas estimaciones por grupos se puede realizar también estimando un modelo denominado multiplicativo (Tabla 4.9 del ANEXO III) que en realidad es una extensión al modelo especificado en la ecuación 2, introduciendo las interacciones de la variable GENERO con las variables explicativas del modelo:

$$\begin{aligned} \ln(\text{Salario}_i) = & \beta_0 + \beta_1 \text{Edu}_i + \beta_2 \text{Exp}_i + \beta_3 \text{Exp}_i^2 + X_i' \lambda \\ & + \delta_0 \text{Genero}_i + \delta_1 (\text{Edu}_i \times \text{Genero}_i) + \delta_2 (\text{Exp}_i \times \text{Genero}_i) + \delta_3 (\text{Exp}_i^2 \times \text{Genero}_i) + (X_i' \text{Genero}_i) \gamma + u_i \end{aligned}$$

Es decir, todas las variables son exactamente las mismas pero se añade un componente referente a las interacciones, permitiendo que todos los coeficientes varíen según el género del trabajador.

Comenzaremos con el análisis del dato más relevante de este estudio que es la brecha de género. Lo que se pretendía con este análisis era comparar hombres y mujeres lo más similares posibles y demostrar que, aun así, las mujeres cobran menos salario que los hombres. Por lo tanto, estamos ante una situación de discriminación salarial por razón de género, ya que individuos con características similares reciben menos salario por el hecho de ser un sexo u otro. En la tabla 4.9 y 4.10 podemos observar la diferencia que hay entre el salario que recibe una mujer en comparación con el del hombre en España.

**Tabla 4.9.Brecha Salarial entre hombres y mujeres por nivel de estudio y grupos de edad.**

DESTU_2	<b>no significativa</b>	DANOS2_2	<b>10,63</b>
DESTU_3	<b>no significativa</b>	DANOS2_3	<b>12,82</b>
DESTU_4	<b>4,03</b>	DANOS2_4	<b>13,48</b>
DESTU_5	<b>5,35</b>	DANOS2_5	<b>15,40</b>
DESTU_6	<b>6,53</b>	DANOS2_6	<b>17,59</b>
DESTU_7	<b>4,91</b>		

Nota: los valores en rojo significan que existe discriminación en contra de las mujeres y los valores en azul significan que existe discriminación a favor de las mujeres (en contra de los hombres).

Fuente: elaboración propia

**Tabla 4.10.Brecha Salarial entre hombres y mujeres por regiones y tipo de mercado laboral**

DNUTS1_2	<b>-3,02</b>	DESTRATO2_2	<b>-2,68</b>
DNUTS1_3	<b>-2,65</b>	DESTRATO2_3	<b>-1,61</b>
DNUTS1_4	<b>-4,66</b>	DMERCADO_2	<b>no significativa</b>
DNUTS1_5	<b>-2,90</b>	DMERCADO_3	<b>no significativa</b>
DNUTS1_6	<b>-2,81</b>	DMERCADO_4	<b>no significativa</b>
DNUTS1_7	<b>-7,67</b>	DCONTROL_2	<b>5,47</b>
DTIPOPAIS_2	<b>-3,20</b>	DTIPOCON_2	<b>-0,95</b>
DRESPONSA_1	<b>4,37</b>	DTIPOJOR_2	<b>-3,50</b>

Nota: los valores en rojo significan que existe discriminación en contra de las mujeres y los valores en azul significan que existe discriminación a favor de las mujeres (en contra de los hombres).

Fuente: elaboración propia

Podemos observar que la brecha salarial es negativa en todas las regiones española, lo que significa que las mujeres residentes en el resto de España ganan del orden de 3-5% más que los hombres en comparación con la región de referencia (Noroeste: Galicia, Principado de Asturias y Cantabria). En especial, destacar Islas Canarias donde las mujeres ganan un 3% más que el Noroeste, frente a los hombres que ganan un 4,7% menos que en el Noroeste. Con la realización de la tabla 4.10 podemos

observar que la brecha salarial en las diferentes regiones es negativa en favor de las mujeres. Por lo tanto, podemos concluir que existe menos brecha a niveles regionales.

Al estudiar la variable edad hay que indicar que se divide en 6 grupos y el grupo referente son las personas menores de 19 años. Se observa que la brecha salarial entre hombres y mujeres aumenta conforme aumenta la edad, pasando de un 11% a 17,5%. Es decir, la renta laboral de las mujeres adultas es un 6% mayor que la del grupo de referencia (menores de 19 años), mientras que los hombres adultos reciben un 21%-24% más que el grupo de referencia.

Cuando hacemos referencia a la variable educación, debemos tener en cuenta que la variable de referencia es ESTU1 –variable que nos explica a los individuos sin educación-. Se observa que la brecha salarial entre hombres y mujeres no es significativa para los bajos niveles de educación (primaria y primera etapa de secundaria). Los individuos con un nivel de estudios mayor reciben mayor salario, pero más los hombres que las mujeres, generando una brecha salarial del orden del 4%-6%.

### **4.3. Descomposición de Oaxaca-Blinder**

Otra alternativa para analizar la discriminación salarial entre distintos grupos de trabajadores es a través de la descomposición de las diferencias de sus ingresos, aplicando la metodología de Blinder-Oaxaca (1973). Para ello, en primer lugar, describimos la metodología utilizada y en segundo lugar comentaremos los resultados obtenidos.

La descomposición de Blinder-Oaxaca ha sido ampliamente utilizada en los estudios empíricos de la discriminación en el mercado laboral (Blinder 1973; Oaxaca 1973). Los economistas y sociólogos lo han utilizado, por ejemplo, para descomponer las diferencias salariales en función del género (por ejemplo, Stanley y Jarrell 1998; Weichselbaumer y Winter-Ebmer 2005) y la raza (por ejemplo, Darity, Guilkey y Winfrey 1996; Kim 2010). Si bien las descomposiciones de Blinder-Oaxaca han sido un pilar en la investigación empírica sobre la discriminación, pueden aplicarse, en principio, para explicar las diferencias en cualquier variable continua entre dos grupos. Los investigadores, por ejemplo, lo utilizaron para examinar la asimilación de inmigrantes (LaLonde y Topel 1992), las tasas de matriculación escolar (Borooah e Iyer 2006), la cobertura de seguro de salud (Bustamante, Fang, Rizzo y Ortega 2009), la

prevalencia de fumar (Bauer, Göhlmann, y Sinning 2007), o incluso las tasas locales de arrendamiento de caza (Munn y Hussain 2010).

#### 4.3.1. Metodología

La metodología de este trabajo se va a basar en la descomposición de Oaxaca-Blinder. Este método sirve para determinar a qué dos componentes se debe la diferencia salarial existente entre dos grupos de personas, en nuestro caso, entre hombres y mujeres. El primer componente se refiere a las diferencias que existen en las variables explicativas, es decir, las observables. Este componente representa las diferencias en las características de los individuos tales como la edad, la educación o el sector en el que trabajan. El segundo componente se refiere a las características no observables de ambos grupos medidas por los parámetros, es decir, a la discriminación.

El objetivo es medir cuánto de la diferencia salarial se debe a las diferencias en las variables explicativas y cuánto se debe a otros factores no observados.

Tal como se ha mencionado anteriormente, la diferencia salarial se introduce mediante una variable ficticia:

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_k X_{ki} + \alpha G_i + \varepsilon_i \quad (3)$$

donde  $G$  es la variable ficticia que toma el valor 1 para un grupo, en nuestro caso *Hombres* y el valor 0 para el otro grupo, es decir, *Mujeres*.

Este otro método nos permite diferenciar por grupos la influencia de los regresores sobre la endógena, cosa que no nos lo permitía el método de variable ficticia. El principal objetivo es conocer cuánto de la diferencia en la endógena entre los grupos se debe a las diferencia entre los predictores y cuánto se debe a otros factores:

$$E(Y_m) - E(Y_h) = R \quad (4)$$

Siendo  $E(Y_m)$  la esperanza matemática de la variable salario para el grupo mujeres y  $E(Y_h)$  la esperanza matemática de la variable salario para el grupo hombres. Primeramente se estiman las ecuaciones de ingreso separadamente para el grupo de hombres y de mujeres:

$$\begin{aligned} Y_m &= \beta_m X_m + u_m \\ Y_h &= \beta_h X_h + u_h \end{aligned} \quad (5)$$

Donde los subíndices h y m hacer referencia al grupo de hombres y mujeres respectivamente. La variable Y es el salario mensual por hora, objeto de estudio y la variable X es la matriz de características individuales.

Una vez realizadas las estimaciones, existen dos maneras de expresar la descomposición. La primera, en tres componentes, denominada “*threefold decomposition*” y la otra, de dos componentes, denominada “*twofold decomposition*”. La primera metodología, según *Jann* (2008), expresa la diferencia en la media del salario mensual por hora entre hombres y mujeres de la siguiente manera:

$$\overline{\Delta Y} = \overline{X'_H} \hat{\beta}_H - \overline{X'_M} \hat{\beta}_M \quad (6)$$

Donde  $\overline{\Delta Y}$  es la diferencia de los salarios medios para las observaciones en el Grupo hombre y el Grupo de mujeres. Es decir,  $\overline{\Delta Y} = \bar{Y}_H - \bar{Y}_M$

La expresión anterior se puede expresar como la suma de tres componentes:

$$\overline{\Delta Y} = \underbrace{(\bar{X}_h - \bar{X}_m)' \hat{\beta}_h}_{\text{dif. en dotaciones}} + \underbrace{\bar{X}'_m (\hat{\beta}_h - \hat{\beta}_m)}_{\text{dif. en coeficientes}} + \underbrace{(\bar{X}_h - \bar{X}_m)' ((\hat{\beta}_h - \hat{\beta}_m)}_{\text{interacción}} \quad (7)$$

La expresión anterior se denomina “*threefold decomposition*”.

Donde:

- El primer término  $\underbrace{(\bar{X}_h - \bar{X}_m)' \hat{\beta}_h}_{\text{dif. en dotaciones}}$  representa la parte de la diferencia de salario explicada por la diferencia en las características o dotaciones grupales.
- El segundo componente  $\bar{X}'_m (\hat{\beta}_h - \hat{\beta}_m)$  mide la contribución de la diferencia en los coeficientes asociado con el efecto de la discriminación laboral en contra de las mujeres.
- El último componente,  $(\bar{X}_h - \bar{X}_m)' ((\hat{\beta}_h - \hat{\beta}_m)}$  recoge el hecho de que las diferencias en las características y en los coeficientes existen a la vez.

Otra vía alternativa de descomposición de las diferencias de salarios entre hombres y mujeres es el método de la doble descomposición denominada “*twofold decomposition*”. El doble enfoque descompone la diferencia de resultado promedio con respecto a un vector de coeficientes de referencia  $\hat{\beta}_R$ . En la literatura de investigación sobre discriminación en el mercado laboral, el vector de coeficiente de referencia generalmente se ha interpretado como no discriminatorio, en otras palabras, como el

conjunto de coeficientes de regresión que surgirían en un mundo sin discriminación en el mercado laboral. Dicha descomposición se puede expresar de la siguiente manera:

$$\overline{\Delta Y} = \underbrace{(\overline{X}_h - \overline{X}_m)' \hat{\beta}_R}_{\text{dif. explicada}} + \underbrace{\overline{X}'_h (\hat{\beta}_h - \hat{\beta}_R)}_{\text{no explicada H}} + \underbrace{\overline{X}'_m (\hat{\beta}_m - \hat{\beta}_R)}_{\text{no explicada M}}$$

no explicada=discriminación laboral

De esta última expresión, se observa que el primer componente corresponde a la diferencia en las características entre hombres y mujeres y el segundo componente indica la diferencia que se debe a la discriminación.

Como muestra la Ecuación 8, la doble descomposición divide la diferencia en los salarios medios en una parte que se explica por las diferencias entre grupos en las variables explicativas, y una parte que no se explica por estas diferencias.

La parte no explicada de la brecha salarial media se ha atribuido a menudo a la discriminación, pero también puede ser resultado de la influencia de variables no observadas. Asimismo, la parte no explicada puede descomponerse en dos subcomponentes, conocidos como "inexplicable A" y "inexplicable B". Si uno interpreta que el vector de coeficiente de referencia no es discriminatorio, estos subcomponentes miden la parte de la diferencia media en los resultados que se origina a partir de la discriminación en favor de los hombres (grupo h) y la parte que proviene de la discriminación contra las mujeres (Grupo m), respectivamente.

La elección de los coeficientes de referencia  $\hat{\beta}_R$ , generalmente depende del investigador. En la literatura sobre discriminación en el mercado laboral, a menudo se asume que solo uno de los dos grupos se enfrenta a la discriminación, por ejemplo, que solo las mujeres o los miembros de minorías étnicas están discriminados. En tales casos, los coeficientes de referencia serán simplemente los coeficientes de una regresión en las observaciones de uno de los grupos. Es decir:

$$\hat{\beta}_R = \hat{\beta}_h \quad \text{o} \quad \hat{\beta}_R = \hat{\beta}_m \quad (9)$$

Algunos investigadores en su lugar han utilizado un promedio ponderado de  $\hat{\beta}_h$  y  $\hat{\beta}_m$  para el cálculo de los coeficientes de referencia. Reimers (1983), por ejemplo, propone dar igual peso a los coeficientes de las regresiones en el Grupo A y las observaciones del Grupo B:

$$\hat{\beta}_R = 0,5\hat{\beta}_h + 0,5\hat{\beta}_m \quad (10)$$

Cotton (1988) sugiere ponderar los coeficientes por la proporción de observaciones en el grupo correspondiente:

$$\hat{\beta}_R = \frac{N_h}{N_h + N_m} \hat{\beta}_h + \frac{N_m}{N_h + N_m} \hat{\beta}_m \quad (11)$$

Otros investigadores aún han defendido el uso de estimaciones de coeficientes de una regresión que agrupa las observaciones de los Grupos A y B, e incluye (Jann 2008) o no incluye (Neumark 1988) la variable indicadora de grupo como un regresor adicional. El paquete *oaxaca* estima los resultados de todas las opciones mencionadas de  $\hat{\beta}_R$ , y también permite a los usuarios especificar sus propios pesos personalizados para  $\hat{\beta}_h$  y  $\hat{\beta}_m$  para construir un conjunto de coeficientes de referencia basados en promedios ponderados.

#### 4.3.2. Resultados

Para obtener los resultados se ha utilizado el programa informático R dado que dispone de un paquete específico para obtener la descomposición de Blinder-Oaxaca. Dicho paquete se denomina “**R package oaxaca**”. En el ANEXO II se adjuntan los códigos utilizados para obtener todos los resultados que aparecen a continuación. Varias implementaciones de software de la descomposición de Blinder-Oaxaca ya están disponibles. Estos incluyen los módulos *oaxaca* (Jann 2008) y *decomp* (Watson 2010) para Stata (StataCorp 2017) que estiman la descomposición de los modelos de regresión lineal. Además, los módulos Stata *fairlie* (Jann 2006) y *nldecompose* (Sinning, Hahn y Bauer 2008) implementan la descomposición de una gran variedad de modelos no lineales utilizando los métodos propuestos en Fairlie (2005), Bauer y Sinning (2008) y Bauer y Sinning (2010). También está disponible una implementación SAS (SAS Institute 2017) de la descomposición de Blinder-Oaxaca para modelos no lineales (Fairlie 2013).

El paquete *oaxaca* es el primer paquete de descomposición Blinder-Oaxaca para el lenguaje de programación estadística R (R Core Team 2017). Implementa varios tipos de descomposición para modelos de regresión lineal, y obtiene estimaciones puntuales de todos los componentes de descomposición utilizando los mismos procedimientos de estimación que el módulo Stata *oaxaca* (Jann 2008). Los errores estándar se calculan utilizando un enfoque de bootstrap no paramétrico

(Efron 1979). A diferencia de cualquier otra implementación de software existente de la descomposición de Blinder-Oaxaca, el paquete *oaxaca* permite a los usuarios generar visualizaciones de gráficos de barras de todos los resultados de descomposición, lo cual facilita la interpretación de los resultados.

Los primeros resultados de la descomposición muestran que la diferencia entre lo que cobran los hombres y lo que cobran las mujeres es aproximadamente un 17,87% favorable a los hombres. (Tabla 4.11)

**Tabla 4.11. Resultados de la descomposición.**

$\overline{\Delta Y_h} = 2.458$	$\overline{\Delta Y_m} = 2.280$	$\overline{\Delta Y_h} - \overline{\Delta Y_m} = 0.178$
---------------------------------	---------------------------------	---

Fuente: Elaboración propia a partir de la estimación realizada con el paquete Oaxaca de R

Los resultados de la **descomposición triple** (tabla 4.12) sugieren que, de esa diferencia de 17,87%, se puede descomponer en tres partes: aproximadamente un -0,17% puede deberse a las características personales de los individuos; un 14% a la diferencia en los coeficientes, es decir, a la discriminación y el resto, 4,04% se explica por la interacción de los dos. El componente atribuido a las diferencias en dotaciones es negativo y significativo, lo que sugiere que las dotaciones superiores de las mujeres respecto al grupo de hombres contribuyen a disminuir la brecha salarial entre estos dos grupos.

**Tabla 4.12. Resultados de triple descomposición**

Resultados de triple descomposición		
coef(endowments)	coef(coefficients)	coef(interaction)
-0.00173 (0,0017)	0.1400 (0,0023)	0.0404 (0,00158)

Nota: Los valores entre paréntesis son las desviaciones típicas estimadas.

Fuente: Elaboración propia a partir de la estimación realizada con el paquete Oaxaca de R

A continuación se analizan los resultados más destacados variable por variable a partir de gráficos que representan la descomposición de la diferencia en el salario en los tres componentes: diferencia en las características, discriminación salarial e interacción entre ambas. Centrándose en la discriminación salarial. En todos los gráficos que se observan a continuación hay que tener en cuenta varios aspectos: i) Todos los valores

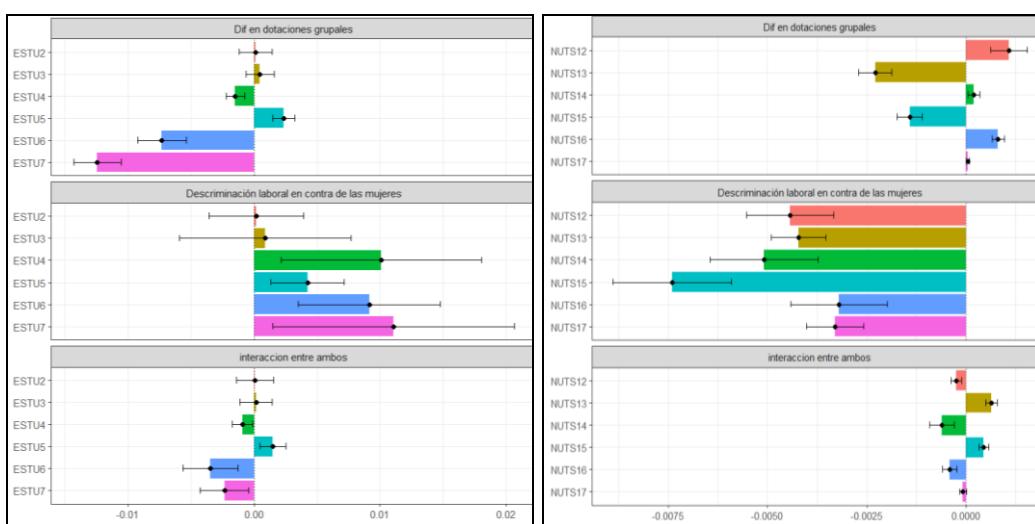
viene acompañados con sus respectivos intervalos de confianzas al 95% de nivel de confianza. ii) En la parte central de cada gráfico se representa la parte de discriminación ( $\bar{X}'_m(\hat{\beta}_h - \hat{\beta}_m)$ , ver ecuación 7) obtenida de la triple descomposición. La interpretación que se hará a continuación se centrará, principalmente, en esta parte del gráfico ya que es el objetivo principal del trabajo. Para facilitar la interpretación, hay que tener en cuenta que un valor positivo indica una discriminación en contra de las mujeres, mientras que un valor negativo indica una discriminación a favor del grupo de las mujeres (o discriminación contra los hombres).

- Para la variable nivel de estudios se observa a partir del gráfico 4.1 (izquierda) la existencia de discriminación significativa contra las mujeres para determinados niveles de estudio. De hecho la discriminación comienza a producirse a partir del nivel 4, es decir, a partir de la segunda etapa de educación secundaria ya que los dos primeros grupos no son significativos. En grupo en el que más se discrimina a las mujeres es el formado por individuos licenciados y doctorados en el que las mujeres reciben un salario un 1 % menor que el de los hombres.
- Localización geográfica: el gráfico 4.1 (derecha) expone que en todas las zonas hay discriminación pero en este caso a favor de las mujeres respecto al grupo de control ya que los valores del componente  $\bar{X}'_m(\hat{\beta}_h - \hat{\beta}_m)$  son negativos y altamente significativos. Para entender la interpretación siempre hay que comparar con la categoría de referencia que en este caso es la región Noroeste: Galicia, Principado de Asturias y Cantabria. Es decir, valores negativos y significativos están indicando que en el resto de las regiones se cobra menos que la región de referencia, pero la brecha de salario entre las regiones es más acusada en caso de los hombres. Cabe destacar que en la zona Este: Castilla y León, Castilla La Mancha y Extremadura; es en la que más se discrimina en favor de las mujeres, sin embargo, no llega al 1%.
- Edad: el gráfico 4.2 (izquierda) manifiesta que en todas las franjas de edad se produce discriminación en contra de las mujeres dado que los valores de dicho componente son positivos y significativos. En los grupos de individuos de 20 a 29 y 30 a 39 años es donde se produce la mayor discriminación ya que las mujeres casi cobran un 5% menos que los hombres. El componente atribuido a las diferencias en dotaciones no es significativo, lo que sugiere que la brecha

salarial entre hombres y mujeres para los diferentes grupos de edades se debe fundamentalmente a la discriminación en contra de las mujeres.

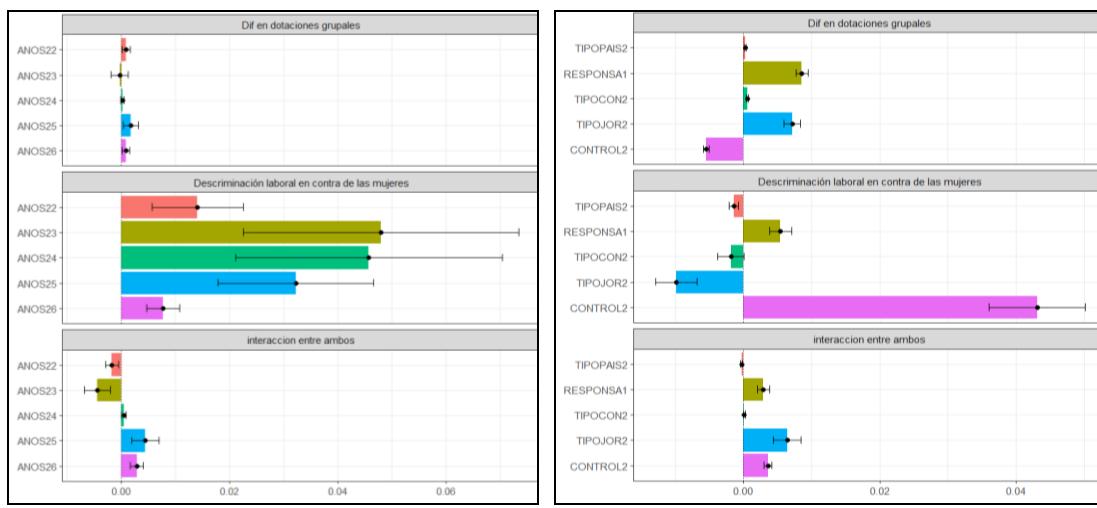
- Resto de variables: Del gráfico 4.2 (derecha) destaca que en, las empresas privadas (CONTROL2), la discriminación en contra de las mujeres se plasma en un salario un 4% menor que el grupo de los hombres. El gráfico 4.3 (izquierda) indica que no existe discriminación según el mercado en el que opere la empresa, pero según el tamaño de la misma sí, existe discriminación a favor de las mujeres. En las empresa con más de 200 trabajadores, las mujeres cobran casi un 2,5 % menos que los hombres respecto a la variable de control.. Por último, el gráfico 4.3 (derecha) muestra que existe discriminación a favor de la mujer en las primeras ocupaciones y en contra de la mujer en el resto, sin embargo, representan un porcentaje de salario muy pequeño.

**Gráfico 4.1. Composición de la diferencia salarial entre hombres y mujeres según el nivel de estudios (izquierda) y según la localización geográfica (derecha).**



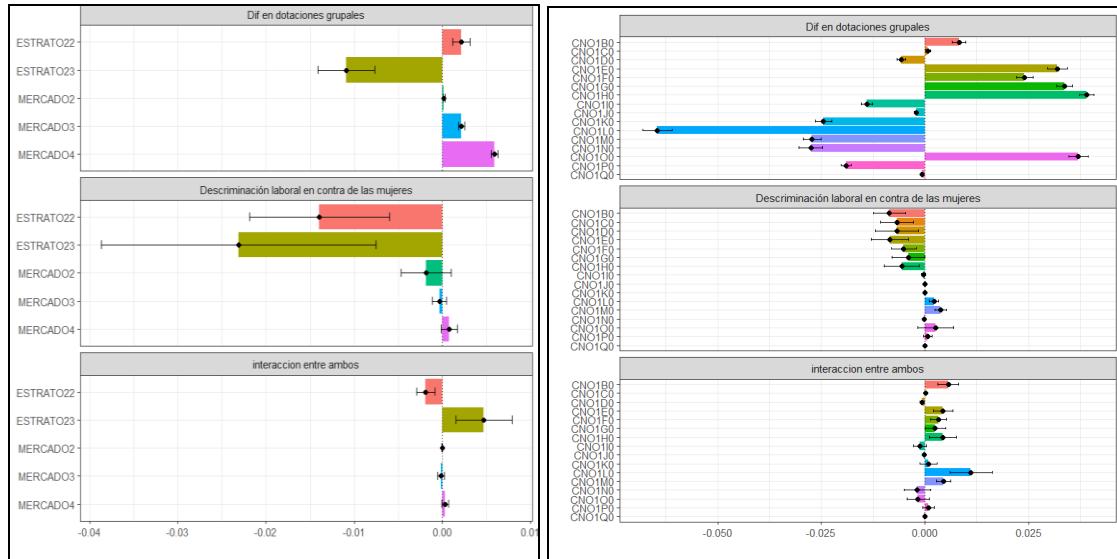
Fuente: Elaboración propia a partir de la estimación realizada con el paquete Oaxaca de R

**Gráfico 4.2. Composición de la diferencia salarial entre hombres y mujeres según la edad (izquierda) y según el resto de variables (derecha).**



Fuente: Elaboración propia a partir de la estimación realizada con el paquete Oaxaca de R

**Gráfico 4.3. Composición de la diferencia salarial entre hombres y mujeres según el resto de variables (izquierda) y según la ocupación (derecha).**



Fuente: Elaboración propia a partir de la estimación realizada con el paquete Oaxaca de R

Para interpretar los resultados de la **descomposición doble** (tabla 4.13), se utiliza la descomposición de Neumark (1988) reflejada con el valor -1 en la columna de “group weight”. Los resultados sugieren que la brecha salarial del 17,87% entre hombres y mujeres es explicada por las diferencias en las características de los individuos en un 5,95 % y el 11,92 % restante, es inexplicable, es decir, es debido a la discriminación. Al contrario que la triple descomposición, el hecho que el componente

atribuido a las diferencias en las dotaciones es positivo y significativo y representa el 35% de las diferencias salariales. Dicho resultado sugiere que las dotaciones inferiores de las mujeres asalariados respecto de los hombres contribuyen a aumentar las diferencias salariales entre estos dos grupos. El componente atribuido a la discriminación salarial es positivo y significativo lo que sugiere la existencia de discriminación en contra de las mujeres, este componente representa el 65% de las diferencias salariales entre ambos grupos. Asimismo, la doble descomposición de Blinde-Oaxaca indica que 5,09% de la parte no explicada se origina a partir de la discriminación en favor de los trabajadores masculinos (componente "inexplicable A"), mientras que 6,82% provienen de la discriminación contra el grupo femenino (componente " inexplicable B "). Los errores estándar proporcionan una idea de la incertidumbre que acompaña a todas las estimaciones puntuales.

**Tabla 3.13. Descomposición de salario por género según la “twofold decomposition”**

$\overline{\Delta Y_h} - \overline{\Delta Y_m} = 0,178$					
Componente debido a diferencias en dotaciones	0,0595 (0,0016)				
Componente atribuido a discriminación	0,119 (0,0018)	Discriminación a favor de hombres (Inexplicable A)	0,00509 (0,0007)	Discriminación en contra de mujeres (Inexplicable B)	0,0682 (0,0018)

*Nota: Entre paréntesis son las desviaciones típicas estimadas (Dt).*

*Fuente: Elaboración propia a partir de la estimación realizada con el paquete Oaxaca de R*

Aunque se ha mostrado la existencia de efectos de las diferencias en las dotaciones grupales y discriminación sobre las brechas de salarios por género, es posible desarrollar una descomposición más detallada de cada uno de estos componentes para que de esta manera observar más claramente donde pueden encontrarse localizados sus factores determinantes. En la Tabla 4.14 se recoge la “twofold decomposition” detallada para cada grupo de variables. Los principales resultados se pueden resumir de la siguiente manera:

- La experiencia, la nacionalidad, responsabilidad en el puesto de trabajo, tipo de contrato, tipo de jornada, tipo de actividad y tipo de mercado contribuyen significativamente al aumento de las diferencias en las dotaciones entre hombres

y mujeres. Sin embargo, el nivel de educación, la localización, el tamaño y tipo de empresa (pública o privada) contrarrestan dichas diferencias.

- Por otra parte, la edad, la educación, la responsabilidad y tipo de empresa son las variables que más contribuyen a explicar los aumentos en la discriminación salarial por género.

**Tabla 4.14. Descomposición detallada de la diferencia salarial por género a partir de la doble descomposición**

	COEFICIENTE DEBIDO A DIFERENCIAS EN DOTACIONES			COEFICIENTE DEBIDO A DISCRIMINACIÓN		
	Coeficiente	Dt	t_ratio	Coeficiente	Dt	t_ratio
CONSTANTE				0,0277	0,0550	0,5041
EXPERIENCIA	0,0143	0,0009	15,5467	0,0157	0,0080	1,9715
EDAD	0,0048	0,0025	1,9374	0,1483	0,0388	3,8240
EDUCACIÓN	-0,0202	0,0021	-9,4998	0,0327	0,0184	1,7723
TIPOPAIS2	0,0002	0,0000	3,7695	-0,0016	0,0004	-4,4423
RESPONSA1	0,0118	0,0004	28,6056	0,0052	0,0010	5,0813
LOCALIZACIÓN	-0,0015	0,0006	-2,4395	-0,0280	0,0034	-8,1758
TIPOCON2	0,0006	0,0001	5,9637	-0,0017	0,0010	-1,7649
TIPOJOR2	0,0150	0,0004	33,5381	-0,0113	0,0010	-10,9702
TIPO DE ACTIVIDAD	0,0371	0,0088	4,2116	-0,0513	0,0186	-2,7576
TAMAÑO DE LA EMPRESA	-0,0085	0,0017	-5,1147	-0,0585	0,0173	-3,3748
CONTROL2	-0,0030	0,0001	-23,6780	0,0442	0,0037	11,8452
TIPO DE MERCADO	0,0091	0,0005	19,7412	-0,0021	0,0024	-0,8530

Nota: Dt son las desviaciones típicas obtenidas a partir del método Bootstrap con el programa R (ver anexo I donde se detallan los códigos utilizados)

Nota: t-ratio es el estadístico t-ratio para contrastar la significatividad individual de cada parámetro estimado-

Nota: los valores en rojo hacen referencia a las variables que contribuyen significativamente al aumento de las diferencias en las dotaciones entre hombres y mujeres y los valores en azul a las variables que más contribuyen a explicar los aumentos en la discriminación salarial por género

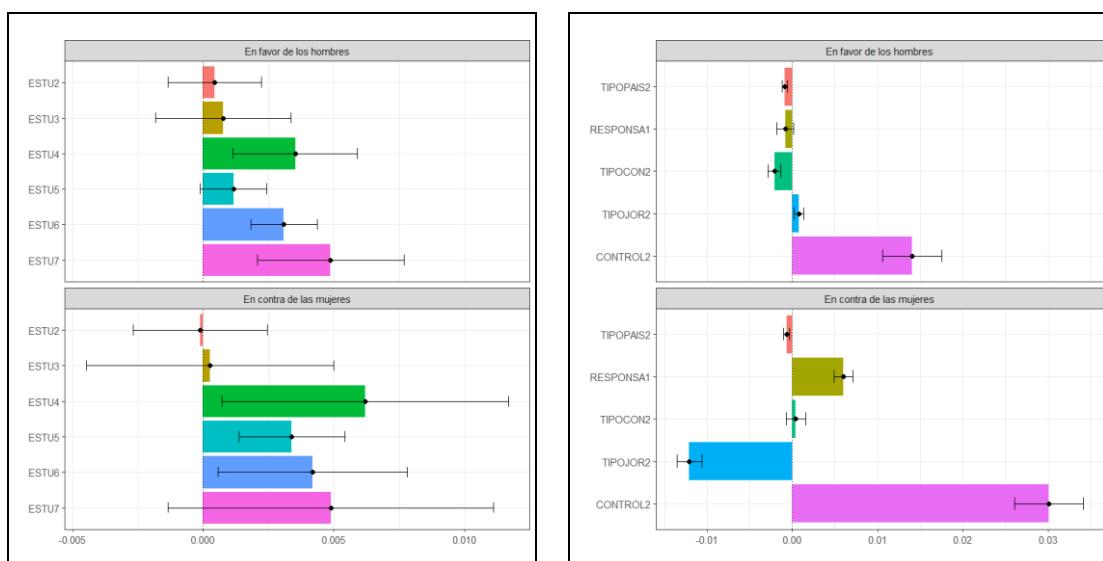
Fuente. Elaboración propia a partir de la estimación realizada con el paquete Oaxaca de R.

Si realizamos un análisis de los resultados más destacados variable por variable a partir de gráficos (ver gráficos 4.4; 4.5 y 4.6 del ANEXO IV), los resultados son los mismos que para la triple descomposición.

Finalmente, para ahondar más en el componente de la discriminación, a través de los siguientes gráficos se puede ver qué proporción de la brecha salarial debido a la discriminación se puede atribuir a la discriminación a favor de los hombres y cual corresponde a la discriminación en contra de las mujeres:

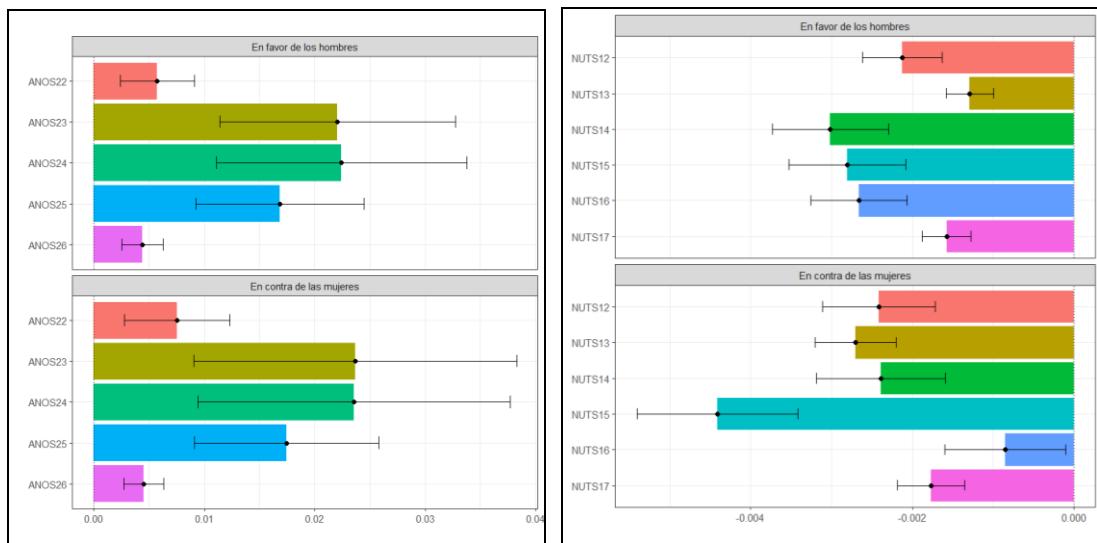
- Nivel de estudios: el gráfico 4.7 (izquierda) indica que para todos los grupos significativos, la discriminación a favor de los hombres es mayor que la discriminación en contra de las mujeres.
- Localización geográfica: el grafico 4.7 (derecha) muestra que en todas las zonas excepto en el Centro y en el Sur, la discriminación en favor de las mujeres es mayor que la discriminación a favor de los hombres.
- Edad: el gráfico 4.8 (izquierda) expone que en todos los grupos de edad, la discriminación en contra de las mujeres es mayor que la discriminación a favor de los hombres pero la diferencia es muy pequeña.
- Resto de variables: en el gráfico 4.8 (derecha) cabe destacar que, tanto según el tipo de jornada y el control de la empresa, la discriminación en contra de las mujeres es aproximadamente un 1.5% mayor que la discriminación en favor de los hombres para ambas variables. Por último, el grafico 4.9 (izquierda) muestra que a medida que la empresa es más grande, la discriminación en contra de las mujeres es cada vez mayor que la discriminación a favor de los hombres; y el gráfico 4.9 (derecha) muestra que en las ocupaciones cualificadas y no manuales, la discriminación en contra de los mujeres es mayor que la discriminación a favor de los hombres.

**Gráfico 4.7. Composición de discriminación según el nivel de estudios (izquierda) y según la localización geográfica (derecha).**



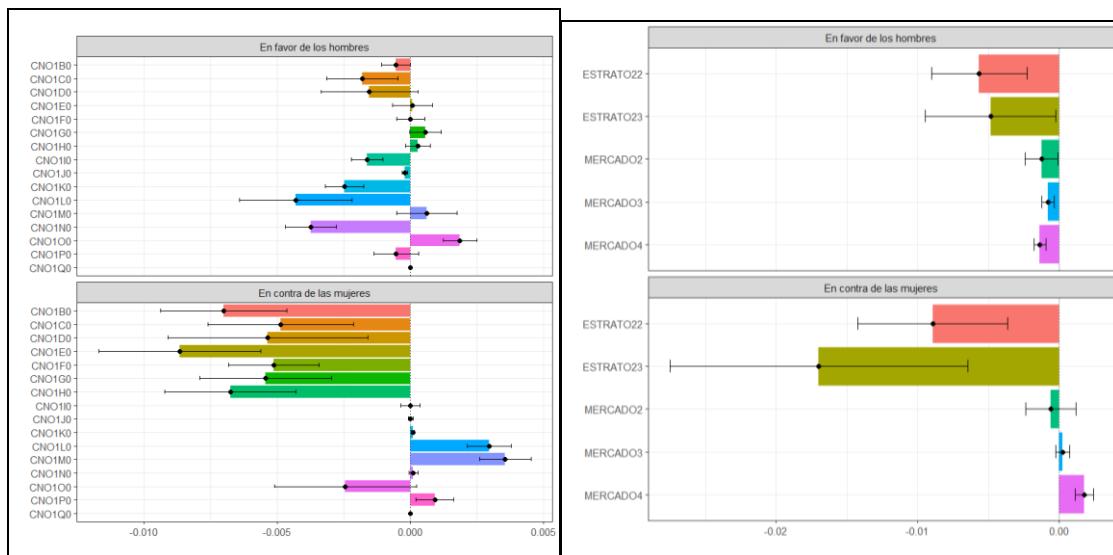
Fuente: Elaboración propia a partir de la estimación realizada con el paquete Oaxaca de R

**Gráfico 4.8. Composición de la discriminación según la edad (izquierda) y según el resto de variables (derecha).**



Fuente: Elaboración propia a partir de la estimación realizada con el paquete Oaxaca de R

**Gráfico 4.9. Composición de la discriminación según el resto de variables (izquierda) y según la ocupación (derecha).**



Fuente: Elaboración propia a partir de la estimación realizada con el paquete Oaxaca de R

## **5. CONCLUSIONES**

Tras haber trabajado con los datos de la Encuesta de Estructura Salarial de 2014, la evidencia empírica demuestra la existencia de discriminación salarial por género en España. Mayoritariamente, la diferencia en el salario entre hombres y mujeres se debe a un componente inexplicable que representa la discriminación, sobretodo, en contra de las mujeres.

A través de modelos econométricos MCO multiplicativos y de la descomposición de Oaxaca-Blinder, el trabajo realizado nos permite afirmar que, en general, las mujeres cobran entre un 15-17% menos que los hombres. Estas diferencias aumentan con los años de experiencia y la edad, el nivel de estudios y el tamaño de la empresa.

Las mayores diferencias en el salario, y por lo tanto la mayor discriminación se da en el grupo de individuos licenciados, doctorados o similares; localizados en el Este de España; con una edad de entre 30 y 39 años aunque en la de 40 a 49 también es elevada; españoles con responsabilidad dentro de la empresa; trabajadores a jornada parcial y con contrato determinado que trabajan en empresas privadas de más de 200 trabajadores. Además en estos grupos, la discriminación es mayoritariamente en contra de las mujeres en lugar de a favor de los hombres.

En base a los datos y resultados, la brecha salarial continúa siendo un problema en España que día a día sufren las mujeres. Una distribución más igualitaria de los salarios entre hombres y mujeres contribuiría a un desarrollo más igualitario y una sociedad más justa para todos en la que nadie se ve perjudicado por sus condiciones. Para llevar esto a cabo, es necesario que toda la sociedad sea consciente del problema y, sobretodo, los partidos políticos que son los encargados de velar por los intereses de la sociedad, proponer las leyes y llevarlas a cabo.

## 6. BIBLIOGRAFÍA

- GRADÍN LAGO, C y DEL RÍO OTERO, C (2009): “Aspectos distributivos de las diferencias salariales por razón de género en España: Un análisis por subgrupos poblacionales”. *Hacienda Pública Española*, Nº 189, págs. 9-46.
- PIEDRESCRITA MURILLO HUERTAS, I y SIMÓN PÉREZ, H (2014): “La Gran Recesión y el diferencial salarial por género en España”. *Hacienda Pública Española*, Nº 208, págs. 39-76.
- GENOVEVA MILLÁN, M; SANTOS PITA, M y PÉREZ NARANJO, L (2015): “Análisis del mercado laboral femenino de España: evolución y factores socioeconómicos determinantes del empleo”. *Papeles de población*, Vol. 21, Nº 84.
- CARVIÑO CUERVA. E (2018): “La desigualdad salarial de género en España en el contexto de la crisis económica y la recuperación“. *Panorama social*, Nº 27, págs. 65-87.
- DE LA RICA, S (2017): “Las Brechas de género en el mercado laboral español y su evolución a lo largo del ciclo de vida”. *Revista de Ciencias y Humanidades*, Nº16, págs. 59-75.
- DE LUCIO, J; DEL VALLE, M y VALERO, M. Determinantes de la Brecha Salarial de Género en España (PDF). Ministerio de Sanidad, Servicios Sociales e Igualdad. Colección EME 2012. [Febrero 2019].
- BRINDUSA, A; CONDE-RUIZ, I y MARRA DE ARTÍÑANO, I. Brechas salariales de Género en España (PDF) FEDEA, marzo 2018 [Febrero 2019].
- HLAVAC, MAREK (2018). oaxaca: Blinder-Oaxaca Decomposition in R. R package version 0.1.4. <https://CRAN.R-project.org/package=oaxaca>
- CAMPS CERVERA, V (2018): “Transparencia y brecha salarial”, *Revista Internacional de Transparencia e Integridad*. Nº6
- TORRESANO MELO, DANIEL IVAN (2009): “ANÁLISIS EMPÍRICO DE LA DISCRIMINACIÓN SALARIAL POR GÉNERO Y ETNIA EN EL ECUADOR EN EL AÑO 2008”.

## ANEXO I

**Tabla 3.1. Clasificación de las ocupaciones**

Clasificación	Subdivisión de grupos de Ocupación	Valor de la variable	Grupos principales de Ocupación CNO- 11
<b>Ocupaciones no manuales muy cualificadas</b>	-Directores y gerentes -Técnicos y profesionales científicos e intelectuales de la salud y la enseñanza -Otros técnicos y profesionales científicos e intelectuales -Técnicos, profesionales de apoyo	A0 B0 C0 D0	Grupos 1, 2 y 3
<b>Ocupaciones no manuales poco cualificadas</b>	-Empleados de oficina que no atienden al público - Empleados de oficina que atienden al público -Trabajadores de los servicios de restauración y comercio -Trabajadores de los servicios de salud y el cuidado de personas -Trabajadores de los servicios de protección y seguridad	E0 F0 G0 H0 I0	Grupos 4 y 5
<b>Ocupaciones manuales cualificadas</b>	- Trabajadores cualificados del sector agrícola, ganadero, forestal y pesquero -Trabajadores cualificados de la construcción, excepto los operadores de máquinas. - Trabajadores cualificados de las industrias manufactureras, excepto operadores de instalaciones y máquinas - Operadores de instalaciones y maquinaria fijas, y montadores - Conductores y operadores de maquinaria móvil.	J0 K0 L0 M0 N0	Grupos 6, 7 y 8
<b>Ocupaciones manuales poco cualificadas</b>	- Trabajadores no cualificados en servicios. -Peones de la agricultura, pesca, construcción, industrias manufactureras y transportes. - Ocupaciones militares	O0 P0 Q0	Grupo 9

Fuente: elaboración propia

**Tabla 3.2. Clasificación de las actividades económicas**

Valor de la variable	Categoría
<b>B0</b>	Industrias extractivas: extracción de antracita, hulla y lignito, extracción de crudo de petróleo y gas natural, extracción de minerales metálicos, otras industrias extractivas, actividades de apoyo a las industrias extractivas
<b>C1</b>	Industria manufacturera: industria de la alimentación, fabricación de bebidas, industria del tabaco, industria textil, confección de prendas de vestir, industria del cuero y del calzado
<b>C2</b>	Industria manufacturera: industria de la madera y del corcho, excepto muebles; cestería y espartería, industria del papel
<b>C3</b>	Industria manufacturera: artes gráficas y reproducción de soportes grabados
<b>C4</b>	Industria manufacturera: coquerías y refino de petróleo, industria química, fabricación de productos farmacéuticos, fabricación de productos de caucho y plásticos
<b>C5</b>	Industria manufacturera: fabricación de otros productos minerales no metálicos
<b>C6</b>	Industria manufacturera: metalurgia; fabricación de productos de hierro, acero y ferroaleaciones, fabricación de productos metálicos, excepto maquinaria y equipo
<b>C7</b>	Industria manufacturera: fabricación de productos informáticos, electrónicos y ópticos, fabricación de material y equipo eléctrico, fabricación de maquinaria y equipo no contemplado en otras partes
<b>C8</b>	Industria manufacturera: fabricación de vehículos de motor, remolques y semirremolques, fabricación de otro material de transporte, fabricación de muebles, otras industrias manufactureras, reparación e instalación de maquinaria y equipo
<b>D0</b>	Suministro de energía eléctrica, gas, vapor y aire acondicionado
<b>E0</b>	Suministro de agua, actividades de saneamiento, gestión de residuos y descontaminación: captación, depuración y distribución de agua, recogida y tratamiento de aguas residuales, recogida, tratamiento y eliminación de residuos; valorización, actividades de descontaminación y otros servicios de gestión de residuos
<b>F0</b>	Construcción: construcción de edificios, ingeniería civil, actividades de construcción especializada
<b>G1</b>	Comercio al por mayor y al por menor; reparación de vehículos de motor y motocicletas: venta y reparación de vehículos de motor y motocicletas, comercio al por mayor e intermediarios del comercio, excepto de vehículos de motor y motocicletas
<b>G2</b>	Comercio al por mayor y al por menor; reparación de vehículos de motor y motocicletas: comercio al por menor, excepto de vehículos de motor y motocicletas
<b>H1</b>	Transporte y almacenamiento: transporte terrestre y por tubería, transporte marítimo y por vías navegables interiores, transporte aéreo

<b>H2</b>	Transporte y almacenamiento: almacenamiento y actividades anexas al transporte, actividades postales y de correos
<b>I0</b>	Hostelería: servicios de alojamiento, servicios de comidas y bebidas
<b>J0</b>	Información y comunicaciones: edición, actividades cinematográficas, de vídeo y de programas de televisión, grabación de sonido y edición musical, actividades de programación y emisión de radio y televisión, telecomunicaciones, programación, consultoría y otras actividades relacionadas con la informática, servicios de información
<b>K0</b>	Actividades financieras y de seguros: servicios financieros, excepto seguros y fondos de pensiones, seguros, reaseguros y fondos de pensiones, excepto seguridad social obligatoria, actividades auxiliares a los servicios financieros y a los seguros
<b>L0</b>	Actividades inmobiliarias
<b>M0</b>	Actividades profesionales, científicas y técnicas: actividades jurídicas y de contabilidad, actividades de las sedes centrales; actividades de consultoría de gestión empresarial, servicios técnicos de arquitectura e ingeniería; ensayos y análisis técnicos, investigación y desarrollo, publicidad y estudios de mercado, otras actividades profesionales, científicas y técnicas, actividades veterinarias.
<b>N0</b>	Actividades administrativas y servicios auxiliares: actividades de alquiler, actividades relacionadas con el empleo, actividades de agencias de viajes, operadores turísticos, servicios de reservas y actividades relacionadas con los mismos, actividades de seguridad e investigación, servicios a edificios y actividades de jardinería, actividades administrativas de oficina y otras actividades auxiliares a las empresas
<b>O0</b>	Administración Pública y defensa; Seguridad Social obligatoria
<b>P0</b>	Educación
<b>Q0</b>	Actividades sanitarias y de servicios sociales: actividades sanitarias, asistencia en establecimientos residenciales, actividades de servicios sociales sin alojamiento
<b>R0</b>	Actividades artísticas, recreativas y de entretenimiento: actividades de creación, artísticas y espectáculos, actividades de bibliotecas, archivos, museos y otras actividades culturales, actividades de juegos de azar y apuestas, actividades deportivas, recreativas y de entretenimiento
<b>S0</b>	Otros servicios: actividades asociativas, reparación de ordenadores, efectos personales y artículos de uso doméstico, otros servicios personales

Fuente: elaboración propia

**Tabla 3.3. Descomposición del salario por hora y por características**

VARIABLES	HOMBRE	MUJER	BRECHA
Por nivel de estudios			
Menos que primaria	9,35	7,53	1,82
Educación primaria	10,11	7,82	2,28
Primera etapa de educación secundaria	10,43	8,46	1,96
Segunda etapa de educación secundaria	12,68	9,63	3,04
Enseñanzas de formación profesional de grado superior y similares	13,45	10,29	3,16
Diplomados universitarios y similares	16,92	13,2	3,71
Licenciados y similares, y doctores universitarios	21,13	15,84	5,26
Edad			
Menores o iguales a 19 años	7,42	7,26	0,16
20 – 29 años	9,19	8,59	0,6
30 – 39 años	12,05	10,61	1,44
40 – 49 años	14,23	11,69	2,53
50 – 59 años	15,84	12,05	3,79
Mayores o iguales a 59 años	16,98	12,17	4,81
Nacionalidad			
España	13,62	11,09	2,53
Resto del Mundo	11,92	9,99	1,93
Unidad territorial			
Noroeste	12,44	9,94	2,49
Noreste	14,31	11,66	2,64
Com. Madrid	16,13	12,46	3,66
Centro	11,79	9,91	1,88
Este	13,77	11,18	2,58
Sur	12,54	10,26	2,27
Canarias	11,43	10,12	1,3
Tipo de jornada			
Completa	13,8	11,91	1,88
Parcial	11,09	8,81	2,28
Tipo de contrato			
Indefinido	14,34	11,46	2,87
Determinado	10,31	9,5	0,8
Ocupación			
Directores y gerentes	30,7	22,74	7,96
Técnicos y profesionales científicos e intelectuales de la salud y la enseñanza	18,94	15,26	3,67
Otros técnicos y profesionales científicos e intelectuales	15,48	12,27	3,21
Técnicos, profesionales de apoyo	11,7	9,67	2,03
Empleados de oficina que no atienden al público	8,48	7,39	1,08
Empleados de oficina que atienden al público	10,69	9	1,68
Trabajadores de los servicios de restauración y comercio	11,4	8,65	2,75

Trabajadores de los servicios de salud y el cuidado de personas	10,62	9,03	1,58
Trabajadores de los servicios de protección y seguridad	9,54	8,48	1,06
Trabajadores cualificados del sector agrícola, ganadero, forestal y pesquero	10,81	9,48	1,33
Trabajadores cualificados de la construcción, excepto los operadores de máquinas.	9,28	7,62	1,66
Trabajadores cualificados de las industrias manufactureras, excepto operadores de instalaciones y máquinas	10	8,36	1,64
Operadores de instalaciones y maquinaria fijas, y montadores	11,66	8,54	3,12
Conductores y operadores de maquinaria móvil.	9,72	8,07	1,64
Trabajadores no cualificados en servicios.	19,38	16,07	3,31
Peones de la agricultura, pesca, construcción, industrias manufactureras y transportes.	10,16	0,56	1,6
Ocupaciones militares	5,9	3,98	1,91
Actividad Económica			
Industrias extractivas	12,4	9,27	3,12
Industria manufacturera: alimentación, bebidas, tabaco, textil, confección de prendas de vestir, cuero y calzado	13,03	10,88	2,14
Industria manufacturera: madera y corcho, excepto muebles; cestería y espartería, industria del papel	14,31	11,57	2,74
Industria manufacturera: artes gráficas y reproducción de soportes grabados	21,64	15,6	6,03
Industria manufacturera: coquerías y refino de petróleo, química, productos farmacéuticos, productos de caucho y plásticos	15,5	12,78	2,71
Industria manufacturera: otros productos minerales no metálicos	11,2	8,8	2,32
Industria manufacturera: metalurgia; productos de hierro, acero y ferroaleaciones, productos metálicos, excepto maquinaria y equipo	12,24	10,2	2,04
Industria manufacturera: productos informáticos, electrónicos y ópticos, material y equipo eléctrico, maquinaria y equipo no contemplado en otras partes	12,61	11,31	1,29
Industria manufacturera: vehículos de motor, remolques y semirremolques, otro material de transporte, muebles, otras industrias manufactureras, reparación e instalación de maquinaria y equipo	10,21	8,7	1,49
Suministro de energía eléctrica, gas, vapor y aire acondicionado	11,48	11,6	-0,11
Suministro de agua, actividades de saneamiento, gestión de residuos y descontaminación	13,65	11,24	2,4
Construcción: construcción de edificios, ingeniería civil, actividades de construcción especializada	14,97	11,08	3,89
Comercio al por mayor y al por menor; reparación de vehículos de motor y motocicletas	16,27	14,24	2,03
Comercio al por mayor y al por menor; reparación de vehículos de motor y motocicletas: comercio al por menor, excepto de vehículos de motor y motocicletas	15,56	13,85	1,7
Transporte y almacenamiento: transporte terrestre y por tubería, transporte marítimo y por vías navegables interiores, transporte aéreo	9,94	8,3	1,63
Transporte y almacenamiento: almacenamiento y actividades anexas al	13,77	12,73	1,03

transporte, actividades postales y de correos			
Hostelería: servicios de alojamiento, servicios de comidas y bebidas	15,19	11,76	3,43
Información y comunicaciones	12,52	10,88	1,63
Actividades financieras y de seguros	12,84	11,12	1,72
Actividades inmobiliarias	11,82	9,78	2,03
Actividades profesionales, científicas y técnicas	10,42	8,33	2,08
Actividades administrativas y servicios auxiliares	15,8	16,73	-0,92
Administración Pública y defensa; Seguridad Social obligatoria	13,5	11,69	1,82
Educación	12,99	10,2	2,79
Actividades sanitarias y de servicios sociales	13,44	11,5	1,94
Actividades artísticas, recreativas y de entretenimiento	23,94	19,02	4,92
Otros servicios	13,82	10,37	3,45
Tipo de mercado			
Local o Regional	11,81	10,55	1,26
Nacional	13,69	11,03	2,66
Unión Europea	14,24	11,24	2,99
Mundial	16,66	12,91	3,74
Tipo de empresa			
Pública	15,62	13,75	1,87
Privada	13,23	10,38	2,84
Tamaño de la empresa			
De 1 a 49 trabajadores	11,4	9,52	1,87
De 50 a 199 trabajadores	13,68	10,96	2,71
Más de 200 trabajadores	15,34	12,02	3,32
Incluye estrato 2 y 3	13,72	10,53	3,18

Fuente: elaboración propia

## **ANEXO II - Programación para elaborar los resultados:**

### **Programa Gretl para elaborar los datos a partir de la Encuesta de Estructura Salarial**

```
# Días del mes de octubre por los que se ha percibido salario:  
DIASMES=DRELABM-DSIESPM2  
  
#Días de relación laboral en el año:  
DIASRELABA=DRELABAM*30.42+DRELABAD  
#(ajustamos IF DIASRELABA >365 THEN DIASRELABA =365)  
  
#Días del año por los que se ha percibido salario:  
DIASANO=DIASRELABA-DSIESPA2-DSIESPA4  
  
# IF SIESPM1='1' THEN DO;  
smpl SIESPM1=1 --restrict  
  
SALMES_M=((365/ DIASANO)*(SALBRUTO - GEXTRA))/12  
SALBASE_M=SALMES_M  
COMSAL_M=0.0*COMSAL  
COMSALTT_M=0.0*COMSAL  
EXTRAORM_M=0.0*COMSAL  
PHEXTRA_M=0.0*COMSAL  
COTIZA_M=0.0635*BASE  
smpl --full  
#Para el cálculo de los componentes del salario mensual:  
#IF SIESPM1='6' THEN DO;  
smpl SIESPM1=6 --restrict  
SALBASE_M=(31/DIASMES)*SALBASE  
COMSAL_M=(31/DIASMES)* COMSAL  
COMSALTT_M=(31/DIASMES)*COMSALTT  
  
#Salario mensual total (incluye pagos extraordinarios y pagos por horas extra):  
SALMES_M=SALBASE_M+COMSAL_M+EXTRAORM+PHEXTRA  
COTIZA_M=COTIZA  
smpl --full  
#Salario mensual ordinario (no incluye pagos extraordinarios):  
SALMOR_M=SALBASE_M+COMSAL_M+PHEXTRA  
  
#Salario mensual neto :  
SALNETO_M=SALMES_M-COTIZA_M-IRPFMES  
  
#Salario anual total (incluye pagos extraordinarios):  
#Para el cálculo de los componentes del salario anual:  
VESP_M=(365/DIASANO)*VESP  
GEXTRA_M=(365/DIASANO)* GEXTRA  
SALANUAL_M=(365/DIASANO)*(SALBRUTO+VESP_M)  
  
#Salario anual ordinario (no incluye pagos extraordinarios ni los pagos en especie):  
SALAOR_M=(365/DIASANO)*(SALBRUTO-GEXTRA_M)  
  
#Horas normales de trabajo en el mes de octubre:  
JMP1_M=(JSP1+JSP2/60)*4.35 + HEXTTRA  
  
#Ganancia por hora normal de trabajo:  
SALHORA_M=SALMES_M/JMP1_M
```

## Programa Gretl para el análisis descriptivo

```

set echo off
set messages off
/*
Listando 172 variables:
 0) const      1) ORDENCCC    2) ORDENTRA     3) NUTS1
 4) CNACE      5) ESTRATO2    6) CONTROL      7) MERCADO
 8) REGULACION  9) SEXO        10) TIPOPAIS    11) CNO1
12) RESPONSA   13) ESTU        14) ANOANTI     15) MESANTI
16) TIPOJOR    17) TIPOCON     18) FIJODISM    19) FIJODISD
20) VAL         21) VAN         22) PUENTES     23) JAP
24) JSP1        25) JSP2        26) HEXTTRA     27) DRELABM
28) SIESPM1    29) DSIESPM1    30) SIESPM2     31) DSIESPM2
32) SALBASE    33) EXTRAORM    34) PHEXTRA     35) COMSAL
36) COMSALTT   37) IRPFMES    38) COTIZA      39) BASE
40) DRELABAM   41) DRELABAD    42) SIESPA1     43) DSIESPA1
44) SIESPA2    45) DSIESPA2    46) SIESPA3     47) DSIESPA3
48) SIESPA4    49) DSIESPA4    50) SALBRUTO    51) GEXTRA
52) VESP        53) ANOS2       54) FACTOTAL    55) D_SIESPA1
56) Dsss_1     57) DIASRELABA  58) DIASANO     59) SALMES_M
60) SALBASE_M   61) COMSAL_Mm   62) SALMES_m    63) SALBASE_m
64) COMSAL_m    65) COMSALTT_m  66) EXTRAORM_m  67) PHEXTRA_m
68) COTIZA_m    69) DIASMES     70) SALMOR      71) SALNETO
72) SALANUAL   73) SALAOR      74) VESP_M      75) GEXTRA_M
76) JMP1_M      77) VESP_m      78) GEXTRA_m    79) JMP1_m
80) SALHORA     81) DNUTS1_2    82) DNUTS1_3    83) DNUTS1_4
84) DNUTS1_5    85) DNUTS1_6    86) DNUTS1_7    87) DCNACE_1
88) DCNACE_2    89) DCNACE_3    90) DCNACE_4    91) DCNACE_5
92) DCNACE_6    93) DCNACE_7    94) DCNACE_8    95) DCNACE_9
96) DCNACE_10   97) DCNACE_11   98) DCNACE_12   99) DCNACE_13
100) DCNACE_14  101) DCNACE_15  102) DCNACE_16  103) DCNACE_17
104) DCNACE_18  105) DCNACE_19  106) DCNACE_20  107) DCNACE_21
108) DCNACE_22  109) DCNACE_23  110) DCNACE_24  111) DCNACE_25
112) DCNACE_26  113) DCNACE_27  114) DESTRATO2_0 115) DESTRATO2_1
116) DESTRATO2_2 117) DESTRATO2_3 118) DESTRATO2_4 119) DCONTROL_1
120) DCONTROL_2 121) SALHORA_m  122) EXP_ANOS   123) DMERCADO_1
124) DMERCADO_2 125) DMERCADO_3 126) DMERCADO_4 127) DREGULACION_1
128) DREGULACION_2 129) DREGULACION_3 130) DREGULACION_4 131)
DREGULACION_5
 132) DSEXO_1    133) DSEXO_6    134) DTIPOPAIS_1 135) DTIPOPAIS_2
 136) DCNO1_1    137) DCNO1_2    138) DCNO1_3    139) DCNO1_4
 140) DCNO1_5    141) DCNO1_6    142) DCNO1_7    143) DCNO1_8
 144) DCNO1_9    145) DCNO1_10   146) DCNO1_11   147) DCNO1_12
 148) DCNO1_13   149) DCNO1_14   150) DCNO1_15   151) DCNO1_16
 152) DCNO1_17   153) DRESPONSA_0 154) DRESPONSA_1 155) DESTU_1
 156) DESTU_2    157) DESTU_3    158) DESTU_4    159) DESTU_5
 160) DESTU_6    161) DESTU_7    162) DTIPOJOR_1 163) DTIPOJOR_2
 164) DTIPOCON_1 165) DTIPOCON_2 166) DANOS2_1   167) DANOS2_2
 168) DANOS2_3   169) DANOS2_4   170) DANOS2_5   171) DANOS2_6
*/
# 1) realizar análisis descriptivo de las variables salario por tipo de variables cualitativas
list v = SALBASE_M SALMES_M SALMOR_M SALNETO_M SALANUAL_M SALAOR_M
/*
summary v --by=SEXO --simple
summary v --by=ESTU --simple
summary v --by=TIPOPAIS --simple
summary v --by=TIPOJOR --simple
summary v --by=MERCADO --simple
summary v --by=CONTROL --simple

```

```

summary v --by=TIPOPAIS --simple
summary v --by=NUTS1 --simple
summary v --by=TIPOCON --simple
summary v --by=ANOS2 --simple
*/
# 2) Realizar análisis descriptivo de la brecha salarial entre hombre y mujeres
# por diferentes características personales
# para ello restringir la muestra por tipo de sexo

list vv = DESTU_1 DESTU_2 DESTU_3 DESTU_4 DESTU_5 DESTU_6 DESTU_7
matrix sem = zeros(7,3)

loop foreach i vv
    smpl SEXO == 1 --restrict
    smpl ESTU == i --restrict
    sem[i,1] = mean(SALHORA_M)
    smpl --full
endloop
loop foreach i vv
    smpl SEXO == 6 --restrict
    smpl ESTU == i --restrict
    sem[i,2] = mean(SALHORA_M)
    sem[i,3] = sem[i,1]-sem[i,2]
    smpl --full
endloop
colnames(sem,"Hombre Mujer Brecha")
rownames(sem, "nada PRI S_1 S_2 FP DIP Dr")
print "===="
print " distribución de SALARIO/hora por tipo de estudio"
print "===="
printf "%10.5g",sem
print "===="
/*
list v2 = DCNACE_1 DCNACE_2 DCNACE_3 DCNACE_4 DCNACE_5\
DCNACE_6 DCNACE_7 DCNACE_8 DCNACE_9 DCNACE_10 DCNACE_11 DCNACE_12
DCNACE_13\
DCNACE_14 DCNACE_15 DCNACE_16 DCNACE_17 DCNACE_18 DCNACE_19 DCNACE_20
DCNACE_21\
DCNACE_22 DCNACE_23 DCNACE_24 DCNACE_25 DCNACE_26 DCNACE_27
matrix sem2 = zeros(27,3)

loop foreach i v2
    smpl SEXO == 1 --restrict
    smpl CNACE== i --restrict
    sem2[i,1] = mean(SALHORA_M)
    smpl --full
endloop
loop foreach i v2
    smpl SEXO == 6 --restrict
    smpl CNACE == i --restrict
    sem2[i,2] = mean(SALHORA_M)
    sem2[i,3] = sem2[i,1]-sem2[i,2]
    smpl --full
endloop
colnames(sem2,"Hombre Mujer Brecha")
rownames(sem2, "1 2 3 4 5 6 7 8 10 11 12 13 14 15 16 17 18 1 20 21 22 23 24 25 26 27")
print "===="
print " distribución de SALARIO/hora por tipo de CÓDIGO ACTIVIDAD ECONOMICA"
print "===="
printf "%10.5g",sem2

```

```

print "====="

matrix sem3 = zeros(17,3)

loop i=1..17
    smpl SEXO == 1 --restrict
    smpl CNO1== i --restrict
    sem3[i,1] = mean(SALHORA_M)
    smpl --full
endloop
loop i=1..17
    smpl SEXO == 6 --restrict
    smpl CNO1 == i --restrict
    sem3[i,2] = mean(SALHORA_M)
    sem3[i,3] = sem3[i,1]-sem3[i,2]
    smpl --full
endloop
colnames(sem3,"Hombre Mujer Brecha" )
rownames(sem3, "1 2 3 4 5 6 7 8 10 11 12 13 14 15 16 17")
print "====="
print " distribución de SALARIO/hora por CODIGO DE OCUPACION"
print "===== "
printf "%10.5g",sem3
print "====="

matrix sem4 = zeros(7,3)

loop i=1..7
    smpl SEXO == 1 --restrict
    smpl NUTS1== i --restrict
    sem4[i,1] = mean(SALHORA_M)
    smpl --full
endloop

loop i=1..7
    smpl SEXO == 6 --restrict
    smpl NUTS1 == i --restrict
    sem4[i,2] = mean(SALHORA_M)
    sem4[i,3] = sem4[i,1]-sem4[i,2]
    smpl --full
endloop
colnames(sem4,"Hombre Mujer Brecha" )
rownames(sem4, "NOROESTE NORESTE MADRID CENTRO ESTE SUR CANARIAS")
print "===== "
print " distribución de SALARIO/hora por LOCALIDAD GEOGRÁFICA"
print "===== "
printf "%10.5g",sem4
print "===== "

# otra POR TIPO DE CONTROL
matrix sem5 = zeros(2,3)

loop i=1..2
    smpl SEXO == 1 --restrict
    smpl CONTROL== i --restrict
    sem5[i,1] = mean(SALHORA_M)
    smpl --full
endloop

```

```

loop i=1..2
    smpl SEXO == 6 --restrict
    smpl CONTROL == i --restrict
    sem5[i,2] = mean(SALHORA_M)
    sem5[i,3] = sem5[i,1]-sem5[i,2]
    smpl --full
endloop
colnames(sem5,"Hombre Mujer Brecha" )
rownames(sem5, "PUBLICO PRIVADO")
print "===="
print " distribución de SALARIO/hora por PUBLICO O PRIVADO"
print "===="
printf "%10.5g",sem5
print "===="

# otra POR TIPO DE MERCADO
matrix sem6 = zeros(4,3)

loop i=1..4
    smpl SEXO == 1 --restrict
    smpl MERCADO == i --restrict
    sem6[i,1] = mean(SALHORA_M)
    smpl --full
endloop

loop i=1..4
    smpl SEXO == 6 --restrict
    smpl MERCADO == i --restrict
    sem6[i,2] = mean(SALHORA_M)
    sem6[i,3] = sem6[i,1]-sem6[i,2]
    smpl --full
endloop
colnames(sem6,"Hombre Mujer Brecha" )
rownames(sem6, "REGIONAL NACIONAL UE MUNDIAL")
print "===="
print " distribución de SALARIO/hora por TIPO DE MERCADO"
print "===="
printf "%10.5g",sem6
print "===="

# otra POR NACIONALIDAD
matrix sem7 = zeros(2,3)

loop i=1..2
    smpl SEXO == 1 --restrict
    smpl TIPOPAIS == i --restrict
    sem7[i,1] = mean(SALHORA_M)
    smpl --full
endloop

loop i=1..2
    smpl SEXO == 6 --restrict
    smpl TIPOPAIS == i --restrict
    sem7[i,2] = mean(SALHORA_M)
    sem7[i,3] = sem7[i,1]-sem7[i,2]
    smpl --full
endloop
colnames(sem7,"Hombre Mujer Brecha" )
rownames(sem7, "ESPAÑA RESTO")
print "===="
print " distribución de SALARIO/hora por NACIONALIDAD"

```

```

print "===="
printf "%10.5g",sem7
print "===="
# otra POR TIPO DE JORNADA
matrix sem8 = zeros(2,3)

loop i=1..2
    smpl SEXO == 1 --restrict
    smpl TIPOJOR == i --restrict
    sem8[i,1] = mean(SALHORA_M)
    smpl --full
endloop

loop i=1..2
    smpl SEXO == 6 --restrict
    smpl TIPOJOR == i --restrict
    sem8[i,2] = mean(SALHORA_M)
    sem8[i,3] = sem8[i,1]-sem8[i,2]
    smpl --full
endloop

colnames(sem8,"Hombre Mujer Brecha")
rownames(sem8, "COMPLETA PARCIAL")
#rowname(sem, "nada primaria secundaria_1 secundaria_2 for_prof diplomado licenciado")
print "===="
print " distribución de SALARIO/hora por tipo de jornada"
print "===="
printf "%10.5g",sem8
print "===="

# otra POR TIPO DE contrato
matrix sem9 = zeros(2,3)

loop i=1..2
    smpl SEXO == 1 --restrict
    smpl TIPOCON == i --restrict
    sem9[i,1] = mean(SALHORA_M)
    smpl --full
endloop

loop i=1..2
    smpl SEXO == 6 --restrict
    smpl TIPOCON == i --restrict
    sem9[i,2] = mean(SALHORA_M)
    sem9[i,3] = sem9[i,1]-sem9[i,2]
    smpl --full
endloop

colnames(sem9,"Hombre Mujer Brecha")
rownames(sem9, "INDEFINIDA DETERMINADA")
print "===="
print " distribución de SALARIO/hora por DURACION DEL CONTRATO"
print "===="
printf "%10.5g",sem9
print "===="
# otra POR TIPO DE EMPRESA SEGUN NUMERO DE TRABAJADORES
matrix sem10 = zeros(4,3)

loop i=1..4
    smpl SEXO == 1 --restrict
    smpl ESTRATO2 == i --restrict
    sem10[i,1] = mean(SALHORA_M)

```

```

smpl --full
endloop

loop i=1..4
    smpl SEXO == 6 --restrict
    smpl ESTRATO2 == i --restrict
    sem10[i,2] = mean(SALHORA_M)
    sem10[i,3] = sem10[i,1]-sem10[i,2]
    smpl --full
endloop
colnames(sem10,"Hombre Mujer Brecha" )
rownames(sem10, "1_49 50_199 200_MAS 50_MAS")
print "===="
print " distribución de SALARIO/hora por TIPO DE EMPRESA"
print "===="
printf "%10.5g",sem10
print "===="

# otra POR edades de las personas
matrix sem11 = zeros(6,3)

loop i=1..6
    smpl SEXO == 1 --restrict
    smpl ANOS2 == i --restrict
    sem11[i,1] = mean(SALHORA_M)
    smpl --full
endloop

loop i=1..6
    smpl SEXO == 6 --restrict
    smpl ANOS2 == i --restrict
    sem11[i,2] = mean(SALHORA_M)
    sem11[i,3] = sem11[i,1]-sem11[i,2]
    smpl --full
endloop
colnames(sem11,"Hombre Mujer Brecha" )
rownames(sem11, "menes19 20_29 30_39 40_49 50_59 mas_60")
print "===="
print " distribución de SALARIO/hora por GRUPO DE EDADES"
print "===="
printf "%10.5g",sem11
print "===="

*/

```

## Programa R para la Descomposición de Oaxaca-Blinder

```
# Leer los datos
load("C:/ANDREA_2019/SALARIO_FINAL_2014.RData")
attach(SALARIO_FINAL_2014)
SALARIO_FINAL_2014$ISALHORA_M <- with(SALARIO_FINAL_2014, log(SALHORA_M))
SALARIO_FINAL_2014$sq_ANOANTI <- with(SALARIO_FINAL_2014, (ANOANTI)^2)
names(SALARIO_FINAL_2014)
# Con la función attach hacemos accesible la base de datos seleccionada:
attach(SALARIO_FINAL_2014)
str(SALARIO_FINAL_2014)
head(SALARIO_FINAL_2014)
summary(ISALHORA_M)
factor(SEXO)
factor(sexo,levels=c("1","M"),labels=c("Hombre","Mujer"))
# CODIFICAR UNA VARIABLE CUALITATIVA
SALARIO_FINAL_2014 <- within(SALARIO_FINAL_2014, {
  DSEXO <- Recode(SEXO, '1=1; 6=0', as.factor=TRUE)
})
library(oaxaca, pos=16)

# modelo 1 con todas las variable
results1 <- oaxaca(formula = ISALHORA_M~ANOANTI + sq_ANOANTI +
  ANOS2 + ESTU + TIPOPAIS + RESPONSA + NUTS1+
  TIPOCON + TIPOJOR + CNO1+ ESTRATO2 +CONTROL + MERCADO|DSEXO_6,
  data =SALARIO_FINAL_2014, R = 1000)

results1$y

# TRIPLE DESCOMPOSICION
results1$threefold$overall
summary(results1$reg$reg.pooled.2)$coefficients["ANOANTI",]
summary(results1$reg$reg.pooled.2)$coefficients["ESTU[T.2]",]
summary(results1$reg$reg.pooled.2)$coefficients["ESTU[T.3]",]
summary(results1$reg$reg.pooled.2)$coefficients["ESTU[T.4]",]
summary(results1$reg$reg.pooled.2)$coefficients["ESTU[T.5]",]
summary(results1$reg$reg.pooled.2)$coefficients["ESTU[T.6]",]
summary(results1$reg$reg.pooled.2)$coefficients["ESTU[T.7]",]
summary(results1$reg$reg.pooled.2)

results1$x$x.mean.diff["ESTU2"]
results1$beta$beta.diff["ESTU2"]
results1$x$x.mean.diff["ANOANTI"]
results1$beta$beta.diff["ANOANTI"]

# GRÁFICO DE LA TRIPLE DESCOMPOSICION
# 1 GRÁFICO GENERAL
plot(results1, components = c("endowments", "coefficients"))
# 2. GRÁFICO POR VARIABLES DESGLOSADO
# 2.1 NIVEL DE ESTUDIO
plot(results1, decomposition="threefold",components = c("endowments", "coefficients", "interaction"),
variables=c("ESTU2", "ESTU3", "ESTU4", "ESTU5", "ESTU6", "ESTU7"),
component.labels=c("endowments" = "Dif en dotaciones grupales",
"coefficients" = "Descriminación laboral en contra de las mujeres", "interaction" = "interaccion entre ambos"))
results1$x$x.mean.diff["ESTU[T.6]"]
results1$x$x.mean.diff["ESTU[T.7]"]
results1$beta$beta.diff["ESTU[T.6]"]
results1$beta$beta.diff["ESTU[T.7]"]
```

```

# 2.2 gráfico para localidad
plot(results1, decomposition="threefold",components = c("endowments","coefficients","interaction"),
variables=c("NUTS1[T.2]","NUTS1[T.3]","NUTS1[T.4]","NUTS1[T.5]","NUTS1[T.6]","NUTS1[T.7]"),
component.labels=c("endowments" = "Dif en dotaciones grupales",
"coefficients" = "Descripción laboral en contra de las mujeres","interaction" = "interaccion entre
ambos"))

# 2.3 GRÁFICO PARA EDADES
plot(results1, decomposition="threefold",components = c("endowments","coefficients","interaction"),
variables=c("ANOS2[T.2]","ANOS2[T.3]","ANOS2[T.4]","ANOS2[T.5]","ANOS2[T.6]"),
component.labels=c("endowments" = "Dif en dotaciones grupales",
"coefficients" = "Descripción laboral en contra de las mujeres","interaction" = "interaccion entre
ambos"))

#2.4 GRAFICO PARA RESTO DE VARIABLES
plot(results1, decomposition="threefold",components = c("endowments","coefficients","interaction"),
variables=c("TIPOPAIS[T.2]","RESPONSA[T.1]","TIPOCON[T.2]","TIPOJOR[T.2]","CONTROL[T.2]",
"),
component.labels=c("endowments" = "Dif en dotaciones grupales",
"coefficients" = "Descripción laboral en contra de las mujeres","interaction" = "interaccion entre
ambos"))

#2.5 GRAFICO PARA TAMAÑO DE EMPRESA Y MMERCADO
plot(results1, decomposition="threefold",components = c("endowments","coefficients","interaction"),
variables=c("ESTRATO2[T.2]","ESTRATO2[T.3]","MERCADO[T.2]","MERCADO[T.3]","MERCAD
O[T.4]"),
component.labels=c("endowments" = "Dif en dotaciones grupales",
"coefficients" = "Descripción laboral en contra de las mujeres","interaction" = "interaccion entre
ambos"))

# 2.6 GRÁFICO POR TIPO DE ACTIVIDAD
plot(results1, decomposition="threefold",components = c("endowments","coefficients","interaction"),
variables=c("CNO1[T.B0]",
"CNO1[T.C0]","CNO1[T.D0]","CNO1[T.E0]","CNO1[T.F0]","CNO1[T.G0]",
"CNO1[T.H0]","CNO1[T.I0]","CNO1[T.J0]","CNO1[T.K0]","CNO1[T.L0]","CNO1[T.M0]",
"CNO1[T.N0]","CNO1[T.O0]","CNO1[T.P0]","CNO1[T.Q0]"),
component.labels=c("endowments" = "Dif en dotaciones grupales",
"coefficients" = "Descripción laboral en contra de las mujeres","interaction" = "interaccion entre
ambos"))

# DOBLE DESCOMPOSICION
# imprimimos los resultados mas importantes
results1$twofold$overall
results1$twofold

# realizar los gráficos
# 1. GRÁFICO GENERAL COMPLEJO
plot(results1,decomposition="twofold",group.weight=-1,unexplained.split=TRUE)
plot(results1,decomposition="twofold",type="variables",group.weight=-1)

# 2. GRÁFICO DE LA DESCOMPOSICIÓN EN "DIFERENCIAS EN CARACTÉRISTICA" Y "
DISCRIMINACIÓN POR GÉNERO"
# 2.1 nivel de estudio
plot(results1,decomposition="twofold",type="variables",group.weight=-1,
components=c("explained","unexplained"),component.labels=c("explained" =
"Diferencias debidas a características", "unexplained" = "Discriminación laboral por género"),
variables=c("ESTU[T.2]","ESTU[T.3]","ESTU[T.4]","ESTU[T.5]","ESTU[T.6]","ESTU[T.7]"))
# 2.2

```

```

plot(results1,decomposition="twofold",type="variables",group.weight=-1,
components=c("explained","unexplained"),component.labels=c("explained" =
"Diferencias debidas a características", "unexplained" = "Discriminacion laboral por genero"),
variables=c("NUTS1[T.2]","NUTS1[T.3]","NUTS1[T.4]","NUTS1[T.5]","NUTS1[T.6]","NUTS1[T.7]")
))

# 2.3
plot(results1,decomposition="twofold",type="variables",group.weight=-1,
components=c("explained","unexplained"),component.labels=c("explained" =
"Diferencias debidas a características", "unexplained" = "Discriminacion laboral por genero"),
variables=c("ANOS2[T.2]","ANOS2[T.3]","ANOS2[T.4]","ANOS2[T.5]","ANOS2[T.6]"))

# 2.4
plot(results1,decomposition="twofold",type="variables",group.weight=-1,
components=c("explained","unexplained"),component.labels=c("explained" =
"Diferencias debidas a características", "unexplained" = "Discriminacion laboral por genero"),
variables=c("TIPOPAIS[T.2]","RESPONSA[T.1]","TIPOCON[T.2]","TIPOJOR[T.2]","CONTROL[T.2]"
"))

# 2.5
plot(results1,decomposition="twofold",type="variables",group.weight=-1,
components=c("explained","unexplained"),component.labels=c("explained" =
"Diferencias debidas a características", "unexplained" = "Discriminacion laboral por genero"),
variables=c("ESTRATO2[T.2]","ESTRATO2[T.3]","MERCADO[T.2]","MERCADO[T.3]","MERCAD
O[T.4]"))

# 2.6
plot(results1,decomposition="twofold",type="variables",group.weight=-1,
components=c("explained","unexplained"),component.labels=c("explained" =
"Diferencias debidas a características", "unexplained" = "Discriminacion laboral por genero"),
variables=c("CNO1[T.B0]",
"CNO1[T.C0]","CNO1[T.D0]","CNO1[T.E0]","CNO1[T.F0]","CNO1[T.G0]",
"CNO1[T.H0]","CNO1[T.I0]","CNO1[T.J0]","CNO1[T.K0]","CNO1[T.L0]","CNO1[T.M0]",
"CNO1[T.N0]","CNO1[T.O0]","CNO1[T.P0]","CNO1[T.Q0]"))

# 3. gráfico de descriminacion en favor y contra
# 3.1 grafico para niveles de estudio

plot(results1,decomposition="twofold",group.weight=-1,unexplained.split=TRUE,
components=c("unexplained A","unexplained B"),component.labels=c("unexplained A" =
"En favor de los hombres", "unexplained B" = "En contra de las mujeres"),
variables=c("ESTU[T.2]","ESTU[T.3]","ESTU[T.4]","ESTU[T.5]","ESTU[T.6]","ESTU[T.7]"))

# 3.2 gráfico para localidad
plot(results1,decomposition="twofold",group.weight=-1,unexplained.split=TRUE,
components=c("unexplained A","unexplained B"),component.labels=c("unexplained A" =
"En favor de los hombres", "unexplained B" = "En contra de las mujeres"),
variables=c("NUTS1[T.2]","NUTS1[T.3]","NUTS1[T.4]","NUTS1[T.5]","NUTS1[T.6]","NUTS1[T.7]"
))

# 3.3 GRÁFICO PARA EDADES
plot(results1,decomposition="twofold",group.weight=-1,unexplained.split=TRUE,
components=c("unexplained A","unexplained B"),component.labels=c("unexplained A" =
"En favor de los hombres", "unexplained B" = "En contra de las mujeres"),
variables=c("ANOS2[T.2]","ANOS2[T.3]","ANOS2[T.4]","ANOS2[T.5]","ANOS2[T.6]"))

# 3.4 GRAFICO PARA RESTO DE VARIABLES
plot(results1,decomposition="twofold",group.weight=-1,unexplained.split=TRUE,
components=c("unexplained A","unexplained B"),component.labels=c("unexplained A" =
"En favor de los hombres", "unexplained B" = "En contra de las mujeres"),

```

```

variables=c("TIPOPAIS[T.2]","RESPONSA[T.1]","TIPOCON[T.2]","TIPOJOR[T.2]","CONTROL[T.2]
"))

# 3.5 GRAFICO PARA TAMAÑO DE EMPRESA Y MMERCADO
plot(results1,decomposition="twofold",group.weight=-1,unexplained.split=TRUE,
components=c("unexplained A","unexplained B"),component.labels=c("unexplained A" =
"En favor de los hombres", "unexplained B" = "En contra de las mujeres"),
variables=c("ESTRATO2[T.2]","ESTRATO2[T.3]","MERCADO[T.2]","MERCADO[T.3]","MERCAD
O[T.4]"))

# 3.6 FGRÁFICO POR TIPO DE ACTIVIDAD
plot(results1,decomposition="twofold",group.weight=-1,unexplained.split=TRUE,
components=c("unexplained A","unexplained B"),component.labels=c("unexplained A" =
"En favor de los hombres", "unexplained B" = "En contra de las mujeres"),
variables=c("CNO1[T.B0]", "CNO1[T.C0]", "CNO1[T.D0]", "CNO1[T.E0]", "CNO1[T.F0]", "CNO1[T.G0]",
"CNO1[T.H0]", "CNO1[T.I0]", "CNO1[T.J0]", "CNO1[T.K0]", "CNO1[T.L0]", "CNO1[T.M0]", "CNO1[T.N0]", "CNO1[T.O0]", "CNO1[T.P0]", "CNO1[T.Q0]"))

plot(results1,decomposition="twofold",group.weight=-
1,unexplained.split=TRUE,components=c("unexplained A",
"unexplained B"),component.labels=c("unexplained A" = "en favor de los hombres",
"unexplained B" = "en contra de los las mujeres"),
variables=c("ANOANTI", "sq_ANOANTI"),variable.labels=c("ANOANTI"="superiores",
"sq_ANOANTI"="extranjero"))

```

## ANEXO III – Estimaciones variable ficticia

**Tabla 4.3. Estimación MCO de la ecuación 2**

Modelo TOTAL:MCO, usando las observaciones 1-209436					
Variable dependiente: l_SALHORA_M					
Desviaciones típicas robustas ante heterocedasticidad, variante HC1					
	<i>Coeficiente</i>	<i>Desv. Típica</i>	<i>Estadístico t</i>	<i>valor p</i>	
const	2.23719	0.0262220	85.32	<0.0001	***
ANOANTI	0.0139025	0.000284750	48.82	<0.0001	***
sq_ANOANTI	-8.88084e-05	8.09429e-06	-10.97	<0.0001	***
DESTU_2	0.00571718	0.00648458	0.8817	0.3780	
DESTU_3	0.0186449	0.00640911	2.909	0.0036	***
DESTU_4	0.104323	0.00655412	15.92	<0.0001	***
DESTU_5	0.135569	0.00687707	19.71	<0.0001	***
DESTU_6	0.192265	0.00717722	26.79	<0.0001	***
DESTU_7	0.308760	0.00714047	43.24	<0.0001	***
DANOS2_2	0.0150720	0.0241504	0.6241	0.5326	
DANOS2_3	0.0982902	0.0241055	4.078	<0.0001	***
DANOS2_4	0.155105	0.0241235	6.430	<0.0001	***
DANOS2_5	0.172158	0.0241704	7.123	<0.0001	***
DANOS2_6	0.170582	0.0244908	6.965	<0.0001	***
DNUTS1_2	0.107400	0.00286653	37.47	<0.0001	***
DNUTS1_3	0.0707586	0.00306636	23.08	<0.0001	***
DNUTS1_4	-0.0104796	0.00297552	-3.522	0.0004	***
DNUTS1_5	0.0651646	0.00260425	25.02	<0.0001	***
DNUTS1_6	0.0337106	0.00296152	11.38	<0.0001	***
DNUTS1_7	0.00825592	0.00421610	1.958	0.0502	*
DTIPOPAIS_2	0.0122976	0.00386536	3.182	0.0015	***
DRESPONSA_1	0.156795	0.00259528	60.42	<0.0001	***
DCNO1_2	-0.265751	0.00669579	-39.69	<0.0001	***
DCNO1_3	-0.368249	0.00667761	-55.15	<0.0001	***
DCNO1_4	-0.529241	0.00698899	-75.72	<0.0001	***
DCNO1_5	-0.592021	0.00761118	-77.78	<0.0001	***
DCNO1_6	-0.552920	0.00743705	-74.35	<0.0001	***
DCNO1_7	-0.463848	0.00714589	-64.91	<0.0001	***
DCNO1_8	-0.552386	0.00840804	-65.70	<0.0001	***
DCNO1_9	-0.475842	0.00776305	-61.30	<0.0001	***
DCNO1_10	-0.444463	0.00770197	-57.71	<0.0001	***
DCNO1_11	-0.579554	0.00750399	-77.23	<0.0001	***
DCNO1_12	-0.570511	0.00739754	-77.12	<0.0001	***
DCNO1_13	-0.503255	0.00739334	-68.07	<0.0001	***
DCNO1_14	-0.569521	0.00730631	-77.95	<0.0001	***
DCNO1_15	-0.174225	0.00741850	-23.49	<0.0001	***
DCNO1_16	-0.551217	0.0152735	-36.09	<0.0001	***
DCNO1_17	-1.15365	0.0680037	-16.96	<0.0001	***
DESTRATO2_2	0.0743441	0.00209567	35.48	<0.0001	***
DESTRATO2_3	0.122237	0.00204791	59.69	<0.0001	***
DMERCADO_2	0.0361528	0.00182593	19.80	<0.0001	***
DMERCADO_3	0.0936725	0.00342474	27.35	<0.0001	***
DMERCADO_4	0.136220	0.00270617	50.34	<0.0001	***
DCONTROL_2	-0.0553688	0.00252036	-21.97	<0.0001	***
DTIPOCON_2	-0.0450498	0.00228494	-19.72	<0.0001	***

DTIPOJOR_2	-0.0528168	0.00239511	-22.05	<0.0001	***
DSEXO_1	0.157260	0.00180980	86.89	<0.0001	***
Media de la vble. dep.	2.382412	D.T. de la vble. dep.		0.494209	
Suma de cuad. residuos	25240.01	D.T. de la regresión		0.347190	
R-cuadrado	0.506577	R-cuadrado corregido		0.506468	
F(46, 209389)	4419.227	Valor p (de F)		0.000000	
Log-verosimilitud	-75594.79	Criterio de Akaike		151283.6	
Criterio de Schwarz	151765.4	Crit. de Hannan-Quinn		151425.1	
Contraste de heterocedasticidad de Breusch-Pagan -					
Hipótesis nula: No hay heterocedasticidad					
Estadístico de contraste: LM = 17559.7					
con valor p = P(Chi-cuadrado(46) > 17559.7) = 0					
Contraste de normalidad de los residuos -					
Hipótesis nula: el error se distribuye normalmente					
Estadístico de contraste: Chi-cuadrado(2) = 28450.2					
con valor p = 0					

Fuente: elaboración propia a partir del programa Gretl

**Tabla 4.4. Estimación MCO para contrastar la permanencia estructural**

Modelo 4_1:MCO, usando las observaciones 1-209436					
Variable dependiente: l_SALHORA_M	Coeficiente	Desv. Típica	Estadístico t	valor p	
const	2.34221	0.0229634	102.0	<0.0001	***
ANOANTI	0.0136158	0.000273744	49.74	<0.0001	***
sq_ANOANTI	-7.22246e-05	7.47674e-06	-9.660	<0.0001	***
DESTU_2	0.00277524	0.00707561	0.3922	0.6949	
DESTU_3	0.0143136	0.00698618	2.049	0.0405	**
DESTU_4	0.0984967	0.00709417	13.88	<0.0001	***
DESTU_5	0.135030	0.00742800	18.18	<0.0001	***
DESTU_6	0.177014	0.00759176	23.32	<0.0001	***
DESTU_7	0.296908	0.00749243	39.63	<0.0001	***
DANOS2_2	7.17846e-05	0.0209855	0.003421	0.9973	
DANOS2_3	0.0785002	0.0209419	3.748	0.0002	***
DANOS2_4	0.137524	0.0209648	6.560	<0.0001	***
DANOS2_5	0.157717	0.0210146	7.505	<0.0001	***
DANOS2_6	0.165243	0.0212575	7.773	<0.0001	***
DNUTS1_2	0.107582	0.00299913	35.87	<0.0001	***
DNUTS1_3	0.0702895	0.00306150	22.96	<0.0001	***
DNUTS1_4	-0.00638613	0.00320274	-1.994	0.0462	**
DNUTS1_5	0.0654306	0.00274205	23.86	<0.0001	***
DNUTS1_6	0.0416894	0.00312022	13.36	<0.0001	***
DNUTS1_7	0.0122179	0.00439222	2.782	0.0054	***
DTIPOPAIS_2	0.0179232	0.00366885	4.885	<0.0001	***
DRESPONSA_1	0.177516	0.00242039	73.34	<0.0001	***
DCNO1_2	-0.267989	0.00501372	-53.45	<0.0001	***
DCNO1_3	-0.369004	0.00499893	-73.82	<0.0001	***
DCNO1_4	-0.557529	0.00552455	-100.9	<0.0001	***
DCNO1_5	-0.564611	0.00652301	-86.56	<0.0001	***
DCNO1_6	-0.588003	0.00614154	-95.74	<0.0001	***
DCNO1_7	-0.421662	0.00569236	-74.08	<0.0001	***

DCNO1_8	-0.506793	0.00718719	-70.51	<0.0001	***
DCNO1_9	-0.419287	0.00689375	-60.82	<0.0001	***
DCNO1_10	-0.386486	0.00646768	-59.76	<0.0001	***
DCNO1_11	-0.608324	0.00619812	-98.15	<0.0001	***
DCNO1_12	-0.616656	0.00617289	-99.90	<0.0001	***
DCNO1_13	-0.482601	0.00603209	-80.01	<0.0001	***
DCNO1_14	-0.601199	0.00594098	-101.2	<0.0001	***
DCNO1_15	-0.200202	0.00581637	-34.42	<0.0001	***
DCNO1_16	-0.506703	0.01424448	-35.57	<0.0001	***
DCNO1_17	-1.08925	0.0418203	-26.05	<0.0001	***
DESTRATO2_2	0.0726322	0.00213740	33.98	<0.0001	***
DESTRATO2_3	0.117263	0.00204189	57.43	<0.0001	***
DMERCADO_2	0.0387060	0.00188784	20.50	<0.0001	***
DMERCADO_3	0.100644	0.00347005	29.00	<0.0001	***
DMERCADO_4	0.145524	0.00268580	54.18	<0.0001	***
DCONTROL_2	-0.0469252	0.00253882	-18.48	<0.0001	***
DTIPOJOR_2	-0.0831322	0.00226127	-36.76	<0.0001	***
DTIPOCON_2	-0.0396704	0.00229421	-17.29	<0.0001	***
Media de la vble. dep.	2.382412	D.T. de la vble. dep.		0.494209	
Suma de cuad. residuos	26203.10	D.T. de la regresión		0.353752	
R-cuadrado	0.487749	R-cuadrado corregido		0.487639	
F(45, 209390)	4430.542	Valor p (de F)		0.000000	
Log-verosimilitud	-79516.22	Criterio de Akaike		159124.4	
Criterio de Schwarz	159596.0	Crit. de Hannan-Quinn		159263.0	

Contraste de Chow de diferencia estructural con respecto a DSEXO\_1 -

Hipótesis nula: no hay diferencia estructural

Estadístico de contraste:  $F(46, 209344) = 211.541$

con valor p =  $P(F(46, 209344) > 211.541) = 0$

Fuente: elaboración propia a partir del programa Gretl

**Tabla 4.5. Estimación MCO para el grupo de hombres**

	Coeficiente	Desv. Típica	Estadístico t	valor p	
const	2.31484	0.0363123	63.75	<0.0001	***
ANOANTI	0.0152010	0.000384158	39.57	<0.0001	***
sq_ANOANTI	-0.000123205	1.08738e-05	-11.33	<0.0001	***
DESTU_2	0.00543182	0.00804397	0.6753	0.4995	
DESTU_3	0.0173299	0.00794547	2.181	0.0292	**
DESTU_4	0.115213	0.00816276	14.11	<0.0001	***
DESTU_5	0.146593	0.00853290	17.18	<0.0001	***
DESTU_6	0.213950	0.00927663	23.06	<0.0001	***
DESTU_7	0.326032	0.00920756	35.41	<0.0001	***
DANOS2_2	0.0567975	0.0338578	1.678	0.0934	*
DANOS2_3	0.151458	0.0338042	4.480	<0.0001	***
DANOS2_4	0.210423	0.0338314	6.220	<0.0001	***
DANOS2_5	0.235313	0.0338919	6.943	<0.0001	***
DANOS2_6	0.244715	0.0342716	7.140	<0.0001	***
DNUTS1_2	0.0931770	0.00384901	24.21	<0.0001	***
DNUTS1_3	0.0621564	0.00428247	14.51	<0.0001	***

DNUTS1_4	-0.0302718	0.00398611	-7.594	<0.0001	***
DNUTS1_5	0.0543879	0.00353567	15.38	<0.0001	***
DNUTS1_6	0.0228886	0.00394389	5.804	<0.0001	***
DNUTS1_7	-0.0165406	0.00565034	-2.927	0.0034	***
DTIPOPAIS_2	0.000876710	0.00519459	0.1688	0.8660	
DRESPONSA_1	0.173201	0.00331993	52.17	<0.0001	***
DCNO1_2	-0.284376	0.00848926	-33.50	<0.0001	***
DCNO1_3	-0.376933	0.00844757	-44.62	<0.0001	***
DCNO1_4	-0.555675	0.00920235	-60.38	<0.0001	***
DCNO1_5	-0.573882	0.00958150	-59.89	<0.0001	***
DCNO1_6	-0.586602	0.0106362	-55.15	<0.0001	***
DCNO1_7	-0.451372	0.00893381	-50.52	<0.0001	***
DCNO1_8	-0.551886	0.0102215	-53.99	<0.0001	***
DCNO1_9	-0.467021	0.00952024	-49.06	<0.0001	***
DCNO1_10	-0.438624	0.00946483	-46.34	<0.0001	***
DCNO1_11	-0.557494	0.0103484	-53.87	<0.0001	***
DCNO1_12	-0.603233	0.0109584	-55.05	<0.0001	***
DCNO1_13	-0.474484	0.00934494	-50.77	<0.0001	***
DCNO1_14	-0.585770	0.00995704	-58.83	<0.0001	***
DCNO1_15	-0.215681	0.0105681	-20.41	<0.0001	***
DCNO1_16	-0.546638	0.0171953	-31.79	<0.0001	***
DCNO1_17	-1.10814	0.0689760	-16.07	<0.0001	***
DESTRATO2_2	0.0749528	0.00276946	27.06	<0.0001	***
DESTRATO2_3	0.128753	0.00283056	45.49	<0.0001	***
DMERCADO_2	0.0367248	0.00254775	14.41	<0.0001	***
DMERCADO_3	0.0909210	0.00444673	20.45	<0.0001	***
DMERCADO_4	0.138482	0.00359264	38.55	<0.0001	***
DCONTROL_2	-0.0319405	0.00369468	-8.645	<0.0001	***
DTIPOCON_2	-0.0495920	0.00320768	-15.46	<0.0001	***
DTIPOJOR_2	-0.0764865	0.00441675	-17.32	<0.0001	***
Media de la vble. dep.	2.458790	D.T. de la vble. dep.		0.501732	
Suma de cuad. residuos	15368.51	D.T. de la regresión		0.358024	
R-cuadrado	0.491001	R-cuadrado corregido		0.490810	
F(45, 119897)	2441.045	Valor p (de F)		0.000000	
Log-verosimilitud	-46968.54	Criterio de Akaike		94029.08	
Criterio de Schwarz	94475.04	Crit. de Hannan-Quinn		94163.32	

Fuente: elaboración propia a partir del programa Gretl

**Tabla 4.6. Estimación MCO para el grupo de mujeres**

Modelo 20:MCO, usando las observaciones 1-89493					
Variable dependiente: l_SALHORA_M					
Desviaciones típicas robustas ante heterocedasticidad, variante HC					
const	2.35578	0.0339716	69.35	<0.0001	***
ANOANTI	0.0115762	0.000415238	27.88	<0.0001	***
sq_ANOANTI	-2.61660e-05	1.17430e-05	-2.228	0.0259	**
DESTU_2	0.00357712	0.0107270	0.3335	0.7388	
DESTU_3	0.0126670	0.0106240	1.192	0.2331	
DESTU_4	0.0715858	0.0108091	6.623	<0.0001	***
DESTU_5	0.0899117	0.0114075	7.882	<0.0001	***

DESTU_6	0.144967	0.0114619	12.65	<0.0001	***
DESTU_7	0.273312	0.0114346	23.90	<0.0001	***
DANOS2_2	-0.0648996	0.0299696	-2.166	0.0304	**
DANOS2_3	0.00762987	0.0299104	0.2551	0.7987	
DANOS2_4	0.0603229	0.0299414	2.015	0.0439	**
DANOS2_5	0.0664619	0.0300240	2.214	0.0269	**
DANOS2_6	0.0533684	0.0305750	1.745	0.0809	*
DNUTS1_2	0.125268	0.00421658	29.71	<0.0001	***
DNUTS1_3	0.0856593	0.00432672	19.80	<0.0001	***
DNUTS1_4	0.0147287	0.00437440	3.367	0.0008	***
DNUTS1_5	0.0824957	0.00377798	21.84	<0.0001	***
DNUTS1_6	0.0489552	0.00441333	11.09	<0.0001	***
DNUTS1_7	0.0442699	0.00624256	7.092	<0.0001	***
DTIPOPAIS_2	0.0333976	0.00565316	5.908	<0.0001	***
DRESPONSA_1	0.128995	0.00417216	30.92	<0.0001	***
DCNO1_2	-0.225053	0.0107576	-20.92	<0.0001	***
DCNO1_3	-0.337853	0.0107813	-31.34	<0.0001	***
DCNO1_4	-0.489475	0.0110096	-44.46	<0.0001	***
DCNO1_5	-0.604499	0.0127736	-47.32	<0.0001	***
DCNO1_6	-0.516895	0.0113338	-45.61	<0.0001	***
DCNO1_7	-0.548362	0.0129161	-42.46	<0.0001	***
DCNO1_8	-0.508009	0.0167619	-30.31	<0.0001	***
DCNO1_9	-0.484727	0.0266174	-18.21	<0.0001	***
DCNO1_10	-0.412126	0.0190394	-21.65	<0.0001	***
DCNO1_11	-0.585007	0.0115679	-50.57	<0.0001	***
DCNO1_12	-0.544576	0.0112811	-48.27	<0.0001	***
DCNO1_13	-0.571171	0.0120234	-47.51	<0.0001	***
DCNO1_14	-0.546895	0.0112846	-48.46	<0.0001	***
DCNO1_15	-0.128641	0.0113245	-11.36	<0.0001	***
DCNO1_16	-0.512988	0.0352622	-14.55	<0.0001	***
DCNO1_17	-1.26424	0.248997	-5.077	<0.0001	***
DESTRATO2_2	0.0734987	0.00317647	23.14	<0.0001	***
DESTRATO2_3	0.114220	0.00294833	38.74	<0.0001	***
DMERCADO_2	0.0389952	0.00258724	15.07	<0.0001	***
DMERCADO_3	0.0950740	0.00531797	17.88	<0.0001	***
DMERCADO_4	0.127720	0.00411297	31.05	<0.0001	***
DCONTROL_2	-0.0828572	0.00343605	-24.11	<0.0001	***
DTIPOCON_2	-0.0420629	0.00319060	-13.18	<0.0001	***
DTIPOJOR_2	-0.0393562	0.00279296	-14.09	<0.0001	***
Media de la vble. dep.	2.280046	D.T. de la vble. dep.		0.464656	
Suma de cuad. residuos	9670.700	D.T. de la regresión		0.328811	
R-cuadrado	0.499492	R-cuadrado corregido		0.499240	
F(45, 89447)	1960.754	Valor p (de F)		0.000000	
Log-verosimilitud	-27421.43	Criterio de Akaike		54934.87	
Criterio de Schwarz	55367.35	Crit. de Hannan-Quinn		55066.77	

Fuente: elaboración propia a partir del programa Gretl

**Tabla 4.7. Modelo MCO multiplicativo**

Modelo 16:MCO, usando las observaciones 1-209436

Variable dependiente: 1\_SALHORA\_M

	<i>Coeficiente</i>	<i>Desv. Típica</i>	<i>EstadÍstico t</i>	<i>valor p</i>	
const	2.45193	0.0225001	109.0	<0.0001	***
FEANOANTI	0.00335607	0.000546160	6.145	<0.0001	***
FEsq_ANOANTI	-9.31829e-05	1.50855e-05	-6.177	<0.0001	***
FEDESTU_2	0.000116615	0.0141435	0.008245	0.9934	
FEDESTU_3	0.00187660	0.0139110	0.1349	0.8927	
FEDESTU_4	0.0402805	0.0140926	2.858	0.0043	***
FEDESTU_5	0.0535065	0.0148370	3.606	0.0003	***
FEDESTU_6	0.0653404	0.0149692	4.365	<0.0001	***
FEDESTU_7	0.0490524	0.0147324	3.330	0.0009	***
FEDANOS2_2	0.106261	0.0184462	5.761	<0.0001	***
FEDANOS2_3	0.128215	0.0180857	7.089	<0.0001	***
FEDANOS2_4	0.134802	0.0181397	7.431	<0.0001	***
FEDANOS2_5	0.153993	0.0182944	8.418	<0.0001	***
FEDANOS2_6	0.175927	0.0194695	9.036	<0.0001	***
FEDNUTS1_2	-0.0302311	0.00592305	-5.104	<0.0001	***
FEDNUTS1_3	-0.0265351	0.00599763	-4.424	<0.0001	***
FEDNUTS1_4	-0.0465564	0.00634522	-7.337	<0.0001	***
FEDNUTS1_5	-0.0290274	0.00538037	-5.395	<0.0001	***
FEDNUTS1_6	-0.0281362	0.00618399	-4.550	<0.0001	***
FEDNUTS1_7	-0.0767430	0.00869181	-8.829	<0.0001	***
FEDTIPOPAIS_2	-0.0320303	0.00734348	-4.362	<0.0001	***
FEDRESPONSA_1	0.0436628	0.00493641	8.845	<0.0001	***
FEDCNO1_2	-0.0592452	0.0101275	-5.850	<0.0001	***
FEDCNO1_3	-0.0398487	0.0100799	-3.953	<0.0001	***
FEDCNO1_4	-0.0666560	0.0110705	-6.021	<0.0001	***
FEDCNO1_5	0.0272745	0.0136257	2.002	0.0453	**
FEDCNO1_6	-0.0723301	0.0125051	-5.784	<0.0001	***
FEDCNO1_7	0.0941328	0.0129777	7.253	<0.0001	***
FEDCNO1_8	-0.0455901	0.0171804	-2.654	0.0080	***
FEDCNO1_9	0.0246523	0.0330016	0.7470	0.4551	
FEDCNO1_10	-0.0299867	0.0207143	-1.448	0.1477	
FEDCNO1_11	0.0244682	0.0124199	1.970	0.0488	**
FEDCNO1_12	-0.0617672	0.0128868	-4.793	<0.0001	***
FEDCNO1_13	0.0927739	0.0124103	7.476	<0.0001	***
FEDCNO1_14	-0.0422650	0.0118933	-3.554	0.0004	***
FEDCNO1_15	-0.0894619	0.0118453	-7.553	<0.0001	***
FEDCNO1_16	-0.0385858	0.0397692	-0.9702	0.3319	
FEDCNO1_17	0.154221	0.204049	0.7558	0.4498	
FEDESTRATO2_1	-0.0267579	0.00403485	-6.632	<0.0001	***
FEDESTRATO2_2	-0.0161268	0.00400203	-4.030	<0.0001	***
FEDMERCADO_2	-0.00587692	0.00371947	-1.580	0.1141	
FEDMERCADO_3	-0.00775874	0.00706684	-1.098	0.2722	
FEDMERCADO_4	0.00588616	0.00541434	1.087	0.2770	
FEDCONTROL_2	0.0546925	0.00496551	11.01	<0.0001	***
FEDTIPOCON_2	-0.00951747	0.00451303	-2.109	0.0350	**
FEDTIPOJOR_2	-0.0350124	0.00465227	-7.526	<0.0001	***
ANOANTI	0.0116438	0.000424237	27.45	<0.0001	***
sq_ANOANTI	-2.70764e-05	1.18840e-05	-2.278	0.0227	**
DESTU_2	0.00403224	0.0118455	0.3404	0.7336	

DESTU_3	0.0137832	0.0116729	1.181	0.2377	
DESTU_4	0.0721350	0.0118261	6.100	<0.0001	***
DESTU_5	0.0905888	0.0124420	7.281	<0.0001	***
DESTU_6	0.144606	0.0123529	11.71	<0.0001	***
DESTU_7	0.273525	0.0122276	22.37	<0.0001	***
DANOS2_2	-0.0544714	0.0233468	-2.333	0.0196	**
DANOS2_3	0.0180961	0.0232189	0.7794	0.4358	
DANOS2_4	0.0707146	0.0232499	3.042	0.0024	***
DANOS2_5	0.0770952	0.0233423	3.303	0.0010	***
DANOS2_6	0.0642894	0.0239392	2.686	0.0072	***
DNUTS1_2	0.127542	0.00453064	28.15	<0.0001	***
DNUTS1_3	0.0893023	0.00449880	19.85	<0.0001	***
DNUTS1_4	0.0146530	0.00488645	2.999	0.0027	***
DNUTS1_5	0.0871250	0.00408971	21.30	<0.0001	***
DNUTS1_6	0.0522921	0.00476622	10.97	<0.0001	***
DNUTS1_7	0.0296263	0.00661532	4.478	<0.0001	***
DTIPOPAIS_2	0.0339266	0.00572682	5.924	<0.0001	***
DRESPONSA_1	0.129223	0.00393912	32.81	<0.0001	***
DCNO1_2	-0.225078	0.00821014	-27.41	<0.0001	***
DCNO1_3	-0.338366	0.00821751	-41.18	<0.0001	***
DCNO1_4	-0.489984	0.00862406	-56.82	<0.0001	***
DCNO1_5	-0.603513	0.0113728	-53.07	<0.0001	***
DCNO1_6	-0.515483	0.00912747	-56.48	<0.0001	***
DCNO1_7	-0.549257	0.0112331	-48.90	<0.0001	***
DCNO1_8	-0.506667	0.0151856	-33.36	<0.0001	***
DCNO1_9	-0.493488	0.0321294	-15.36	<0.0001	***
DCNO1_10	-0.410434	0.0194185	-21.14	<0.0001	***
DCNO1_11	-0.583925	0.00940388	-62.09	<0.0001	***
DCNO1_12	-0.543754	0.00911085	-59.68	<0.0001	***
DCNO1_13	-0.572301	0.0102343	-55.92	<0.0001	***
DCNO1_14	-0.544077	0.00899394	-60.49	<0.0001	***
DCNO1_15	-0.126321	0.00880321	-14.35	<0.0001	***
DCNO1_16	-0.510113	0.0367793	-13.87	<0.0001	***
DCNO1_17	-1.26270	0.199734	-6.322	<0.0001	***
DESTRATO2_2	-0.113389	0.00306439	-37.00	<0.0001	***
DESTRATO2_3	-0.0366852	0.00304738	-12.04	<0.0001	***
DMERCADO_2	0.0402487	0.00277281	14.52	<0.0001	***
DMERCADO_3	0.0974463	0.00563248	17.30	<0.0001	***
DMERCADO_4	0.128454	0.00423170	30.36	<0.0001	***
DCONTROL_2	-0.0840872	0.00358136	-23.48	<0.0001	***
DTIPOCON_2	-0.0410922	0.00336347	-12.22	<0.0001	***
DTIPOJOR_2	-0.0393870	0.00289771	-13.59	<0.0001	***

Media de la vble. dep.	2.382412	D.T. de la vble. dep.	0.494209
Suma de cuad. residuos	24985.26	D.T. de la regresión	0.345470
R-cuadrado	0.511557	R-cuadrado corregido	0.511347
F(90, 209345)	2436.127	Valor p (de F)	0.000000
Log-verosimilitud	-74532.50	Criterio de Akaike	149247.0
Criterio de Schwarz	150179.9	Crit. de Hannan-Quinn	149521.0

Contraste de heterocedasticidad de Breusch-Pagan -

Hipótesis nula: No hay heterocedasticidad

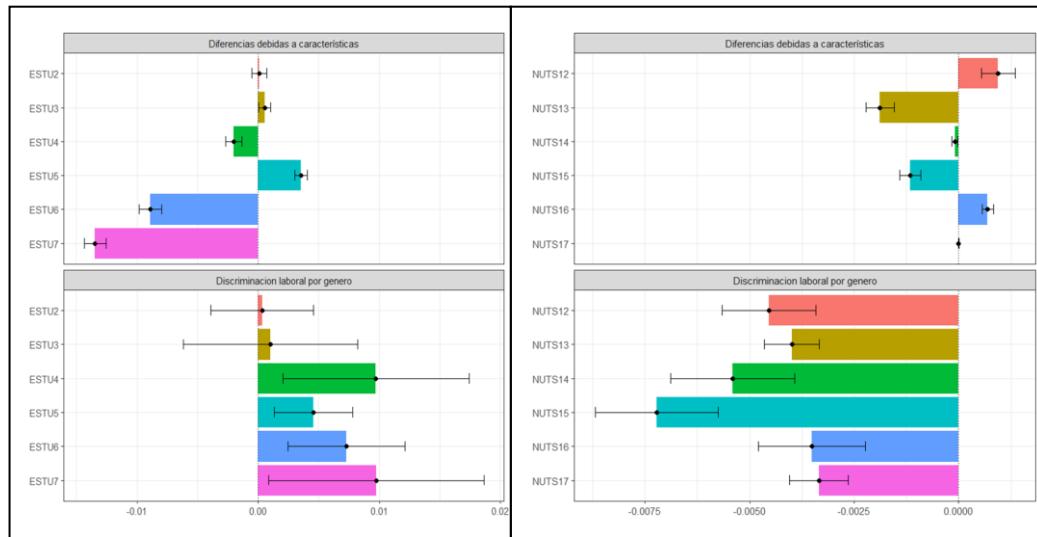
Estadístico de contraste: LM = 18727.7

con valor p = P(Chi-cuadrado(90) > 18727.7) = 0

Fuente: elaboración propia a partir del programa Gretl

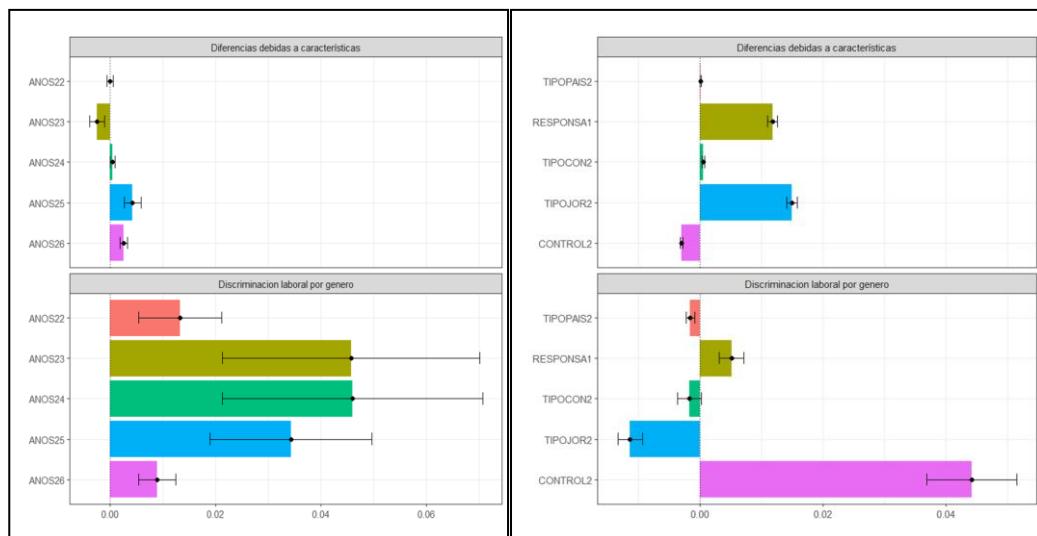
## ANEXO IV – Gráficos descomposición doble

**Gráfico 4.4. Composición de la diferencia salarial entre hombres y mujeres según el nivel de estudios (izquierda) y según la localización geográfica (derecha).**



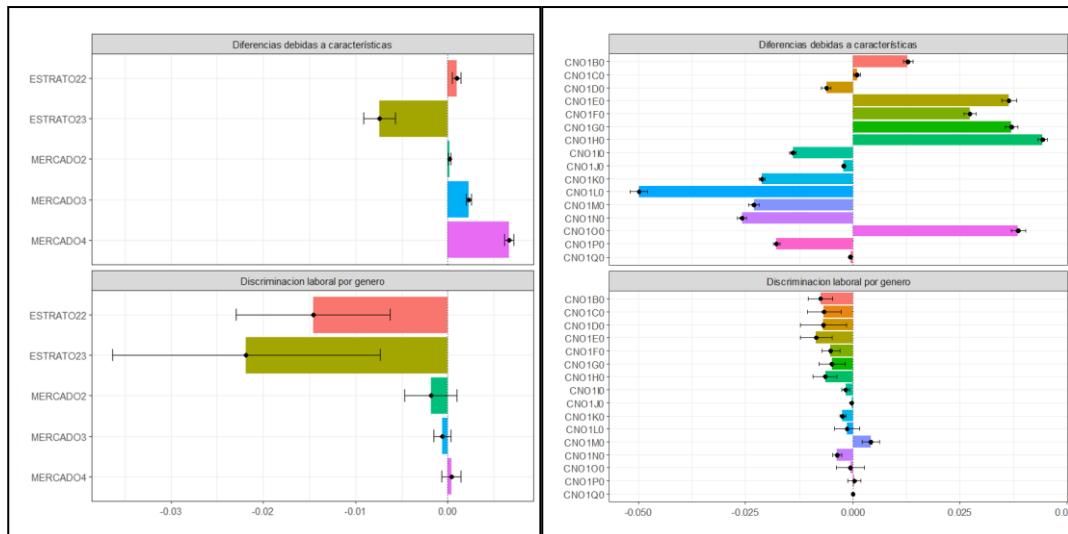
Fuente: Elaboración propia a partir de la estimación realizada con el paquete Oaxaca de R

**Gráfico 4.5. Composición de la diferencia salarial entre hombres y mujeres según la edad (izquierda) y según el resto de variables (derecha).**



Fuente: Elaboración propia a partir de la estimación realizada con el paquete Oaxaca de R

**Gráfico 4.6. Composición de la diferencia salarial entre hombres y mujeres según el resto de variables (izquierda) y según la ocupación (derecha).**



Fuente: Elaboración propia a partir de la estimación realizada con el paquete Oaxaca de R