

Discriminación salarial, oferta salarial y salarios de reserva en España

María Arrazola

e-mail: maria.arezola@urjc.es

teléfono: 914887860

José de Hevia

e-mail: jose.dehevia@urjc.es

teléfono: 914887573

Departamento de Fundamentos del Análisis Económico

Universidad Rey Juan Carlos

Campus de Vicálvaro

Paseo Artilleros s/n 28032 Madrid

Resumen

En este trabajo se proponen diferentes medidas para analizar las diferencias salariales por género tanto en términos de salarios ofertados o potenciales, como de salarios de reserva y de salarios observados. Además se estudia el alcance de dichas medidas en términos del análisis de discriminación laboral de acuerdo con la propuesta de Oaxaca (1973) y se efectúa una aplicación al caso de las diferencias salariales por género en España.

Clasificación JEL: J31

Palabras clave: diferencias salariales, discriminación salarial, oferta salarial, salarios de reserva

1. Introducción

En el contexto del análisis de las diferencias salariales entre diferentes grupos socioeconómicos, por ejemplo entre hombres y mujeres, Oaxaca (1973) y Blinder (1973) propusieron una descomposición de dichas diferencias en dos componentes, uno relacionado con las diferencias en las características de los individuos, y otro relacionado con diferencias en los efectos sobre los salarios de esas características, este último componente es el que se emplea habitualmente como medida de discriminación salarial en el mercado de trabajo. A partir de esos trabajos pioneros, se han sucedido innumerables trabajos aplicados que estudian la discriminación salarial por diferentes motivos (para una revisión de la literatura, véase Altonji y Blank, 1999).

A este respecto, hay que señalar que la literatura aplicada sobre discriminación salarial se ha centrado casi exclusivamente en analizar las diferencias salariales empleando para ello los salarios medios observados de los diferentes subgrupos sociales analizados, por ejemplo hombres y mujeres, es decir, empleando únicamente los salarios de los individuos que trabajan y por tanto, ignorando los salarios ofertados a aquellos individuos que no trabajan. No obstante, recientemente ha existido interés por estudiar hasta qué punto las diferencias salariales observadas reflejan o no las verdaderas diferencias salariales entre subgrupos sociales. Así, por ejemplo, para el caso de las diferencias salariales por raza, Johnson, Kitamura y Neal (2000) y Neal (2004), y por razón de género, Mulligan y Rubinstein (2005), Blau y Kahn (2006) y Olivetti y Petrongolo (2008) muestran que puede existir una brecha sustancial entre las diferencias salariales observadas y las potenciales, Sin embargo, en estos artículos no se muestra interés en analizar los factores discriminatorios que generan esas diferencias salariales. Nuestro objetivo es precisamente combinar ambos

aspectos: estudiar las diferencias en términos de salarios potenciales y los determinantes de dichas diferencias.

En el contexto de análisis de la discriminación salarial centrarse en analizar exclusivamente los salarios observados y no los ofertados o potenciales tiene, al menos tres inconvenientes. El primero es que es un análisis parcial en el que no se tiene en cuenta a una parte relevante de la población: los parados e inactivos, cuya situación, obviamente, no es ajena a la dinámica salarial. La segunda es que en tanto que en la determinación de si se observa o no un salario ofertado a un individuo interviene decididamente el salario de reserva de ese individuo, las diferencias salariales observadas y el correspondiente análisis de discriminación laboral realizado a partir de ellas, puede estar contaminado por otro tipo de discriminaciones sociales, no estrictamente salariales, que afecten a los salarios de reserva y a sus determinantes y no a la oferta salarial. Y la tercera y última es que, cuando uno se centra sólo en el análisis de los salarios observados, y dado el problema de selección muestral existente, el análisis empírico de las diferencias salariales observadas y la discriminación salarial, se complica sustancialmente, siendo además las interpretaciones de los resultados obtenidos muy discutibles (véase, Neuman and Oaxaca, 2004).

En este contexto el objetivo de este trabajo es triple. Primero, proponer diferentes medidas para el análisis de las diferencias salariales: en términos de salarios ofertados o potenciales, de salarios de reserva y de salarios observados. Segundo, estudiar el alcance de dichas medidas en términos del análisis de discriminación laboral y tercero efectuar una aplicación al caso de las diferencias salariales en España por género.

Además de la presente introducción, el artículo tiene otros tres apartados. En el segundo apartado del trabajo, se presentará un marco de análisis para el estudio de diferentes medidas de discriminación salarial en el contexto de los salarios ofertados y de los observados. En el tercer apartado se presentará una aplicación con datos españoles del marco definido en el segundo apartado. El artículo se cierra con un apartado de conclusiones.

2. El marco de referencia

En este apartado analizamos el marco teórico en el que se desarrolla este estudio que permitirá una sistematización del estudio de las diferencias salariales. Primero estudiaremos el modelo salarial considerado y posteriormente las diferencias salariales entre subpoblaciones.

2.1. Ecuación de oferta salarial

El marco teórico empleado es estándar en la literatura (Gronau, 1974 and Lewis, 1974). El modelo está constituido por las siguientes ecuaciones:

$$\log W^* = X\beta_1 + u_1 \quad (1)$$

$$\log W^R = X\beta_2 + u_2 \quad (2)$$

Donde W^* es el salario ofertado y W^R es el salario de reserva, es decir, el mínimo salario al que un individuo está dispuesto a aceptar un puesto de trabajo. Además, suponemos que

$$\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \sim N \left[0, \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{pmatrix} \right]$$

Los individuos decidirán aceptar la oferta salarial, y por tanto trabajar, si y sólo si $W^* \geq W^R$, en caso contrario, no trabajarán y no se observará el salario ofertado. En este contexto los salarios observados W presentan la siguiente pauta:

$$W = \begin{cases} W^* & \text{si } W^* \geq W^R \rightarrow \log(W^* / W^R) = X\beta_1 - X\beta_2 + \varepsilon \geq 0 \\ - & \text{si } W^* < W^R \rightarrow \log(W^* / W^R) = X\beta_1 - X\beta_2 + \varepsilon < 0 \end{cases} \quad (3)$$

Donde $\varepsilon = u_1 - u_2$ y se asume que $V(\varepsilon) = \sigma^2 = \sigma_1^2 + \sigma_2^2 - \sigma_{12} = 1$, este es un supuesto que se realiza habitualmente en la literatura y que permite identificar los parámetros del modelo.

En el contexto del modelo definido por las ecuaciones (1), (2) y (3) existen varias funciones con interés económico. Efectivamente se pueden considerar las siguientes funciones de interés:

$$E[\log W^* / X] = X\beta_1 \quad (4)$$

$$E[\log W^R / X] = X\beta_2 \quad (5)$$

$$E[\log(W^* / W^R) / X] = X\beta_1 - X\beta_2 = X(\beta_1 - \beta_2) = X\gamma \quad (6)$$

$$E[\log W^* / X, W^* \geq W^R] = X\beta_1 + \sigma_{\varepsilon_1} \frac{\phi(X\gamma)}{\Phi(X\gamma)} \quad (7)$$

Siendo $\phi(\cdot)$ y $\Phi(\cdot)$, respectivamente, la función de densidad y de distribución de una normal estándar y $\sigma_{\varepsilon u_1} = E(\varepsilon u_1) = \sigma_1^2 - \sigma_{12}$.

La ecuación (4) mide el valor esperado para el logaritmo del salario ofertado o potencial, en función de las características socioeconómicas de los individuos, se observe o no finalmente el salario ofertado. La expresión (5) mide el valor esperado del logaritmo del salario de reserva en función de las características socioeconómicas de los individuos. La expresión (6) mide la diferencia entre los valores esperados del logaritmo del salario ofertado y de los salarios de reserva, que es lo que determina el que finalmente se acepte o no la oferta salarial y el que se disponga de trabajo o no. La expresión (7) mide el valor esperado para el salario ofertado para los individuos que efectivamente trabajan.

Aunque en la literatura aplicada la consideración de modelos como el definido por (1), (2) y (3) es muy frecuente, lo que no es habitual es considerar el análisis de todas las funciones relevantes que dicho modelo proporciona¹. Nuestro planteamiento es que, dependiendo de la cuestión económica que se desee analizar habrá que centrarse en estudiar una u otra expresión. Así, por ejemplo, si como ocurre con la gran mayoría de la literatura de los modelos de acumulación de capital humano, se desea estudiar los determinantes de la oferta salarial uno se debe fijar en (4). Sin embargo, si se está interesado en analizar los

¹ Existe alguna excepción como Arrazola y Hevia (2008), en que para caso del análisis del rendimiento económico de la educación muestran el interés de estas expresiones.

determinantes socioeconómicos de los salarios de reserva de los individuos, uno se ha de fijar en (5). Si el interés se encuentra en los determinantes de la observación o no del salario ofertado, nos fijaremos en (6) y, finalmente, si nuestro interés está en estudiar sólo los salarios observados, hemos de fijarnos en (7).

En cualquier caso ha de quedar claro que, a partir de una muestra de individuos, se puede disponer de estimaciones de los parámetros de las ecuaciones (4) a (7). Efectivamente, a partir, por ejemplo, del método bietápico de Heckman se pueden estimar las ecuaciones (6) y (7), la expresión (6) por medio de un modelo probit y la expresión (7) por MCO; y a partir de esas estimaciones, obtener estimaciones de las ecuaciones (4) y (5).

2.2. Análisis de las diferencias salariales

Como nuestro interés está en disponer de diferentes medidas de discriminación salarial para diferentes grupos socioeconómicos de interés, como hombres y mujeres, vamos a analizar las diferencias en los valores esperados de los salarios entre cada uno de esos grupos. En correspondencia con las expresiones (4) a (7), habría al menos cuatro dimensiones de análisis de diferencias salariales medias Hombres-Mujeres. Así, considerando el subíndice M para los hombres y el subíndice F para las mujeres, tendríamos:

$$D^* = E[E[\log W_M^* / X_M]] - E[E[\log W_F^* / X_F]] = E(X_M)\beta_{1M} - E(X_F)\beta_{1F} \quad (8)$$

$$D^R = E[E[\log W_M^R / X_M]] - E[E[\log W_F^R / X_F]] = E(X_M)\beta_{2M} - E(X_F)\beta_{2F} \quad (9)$$

$$\begin{aligned}
D^{*R} &= E[E[\log(W_M^*/W_M^R)/X_M]] - E[E[\log(W_F^*/W_F^R)/X_F]] \\
&= E(X_M)(\beta_{1M} - \beta_{2M}) - E(X_F)(\beta_{1F} - \beta_{2F}) = E(X_M)\gamma_M - E(X_F)\gamma_F \quad (10) \\
&= D^* - D^R
\end{aligned}$$

$$\begin{aligned}
D^O &= E[E[\log W_M^*/X_M, W_M^* \geq W_M^R]] - E[E[\log W_F^*/X_F, W_F^* \geq W_F^R]] \\
&= E\left(X_M/W_M^* \geq W_M^R\right)\beta_{1M} + \sigma_{\varepsilon_M u_{1M}} E\left(\frac{\phi(X_M \gamma_M)}{\Phi(X_M \gamma_M)} / W_M^* \geq W_M^R\right) \\
&\quad - E\left(X_F/W_F^* \geq W_F^R\right)\beta_{1F} - \sigma_{\varepsilon_F u_{1F}} E\left(\frac{\phi(X_F \gamma_F)}{\Phi(X_F \gamma_F)} / W_F^* \geq W_F^R\right) \quad (11)
\end{aligned}$$

D^* mide la diferencia salarial media que existe entre los hombres y las mujeres en términos de los salarios ofertados o potenciales, trabajen o no los individuos. D^R mide la diferencia salarial media que existe en los salarios de reserva de los hombres y de las mujeres. D^{*R} es la diferencia entre D^* y D^R , y se puede considerar una medida de las diferencias medias existentes entre los determinantes de la observación o no de los salarios. D^O mide la diferencia salarial media observada entre hombres y mujeres.

La literatura de análisis de discriminación salarial se ha centrado fundamentalmente en analizar los salarios observados a partir de expresiones como (11) (ver, por ejemplo, Reimers, 1983, Dolton y Makepeace, 1986, De la Rica y Ugidos, 1995 y Hernández y Méndez, 2005). Si embargo, en este artículo reivindicamos el interés económico que tiene el análisis de las diferencias salariales presentadas en las expresiones (8) a (10). Efectivamente, si deseamos disponer de medidas globales de discriminación salarial, parece

conveniente separar lo que es pura discriminación en la oferta salarial, de lo que es cualquier otro tipo de discriminación social no necesariamente salarial que determine diferencias en los salarios de reserva, lo que, obviamente, no ocurre en la expresión (11).

A este respecto, y aplicando la propuesta de Oaxaca (1973) y suponiendo que la estructura no discriminatoria es la de los hombres, se puede reescribir las expresiones (8) a (11) del siguiente modo:

$$D^* = E(X_F)[\beta_{1M} - \beta_{1F}] + [E(X_M) - E(X_F)]\beta_{1M} \quad (12)$$

$$D^R = E(X_F)[\beta_{2M} - \beta_{2F}] + [E(X_M) - E(X_F)]\beta_{2M} \quad (13)$$

$$D^{*R} = E(X_F)[\beta_{1M} - \beta_{2M} - [\beta_{1F} - \beta_{2F}]] + [E(X_M) - E(X_F)](\beta_{1M} - \beta_{2M}) \quad (14)$$

$$\begin{aligned} D^o &= E(X_F / W_M^* \geq W_M^R)[\beta_{1M} - \beta_{1F}] + \\ &+ [E(X_M / W_M^* \geq W_M^R) - E(X_F / W_F^* \geq W_F^R)]\beta_{2M} + \\ &+ \sigma_{\varepsilon_M u_{1M}} E\left(\frac{\phi(X_M \gamma_M)}{\Phi(X_M \gamma_M)} / W_M^* \geq W_M^R\right) - \sigma_{\varepsilon_F u_{1F}} E\left(\frac{\phi(X_F \gamma_F)}{\Phi(X_F \gamma_F)} / W_F^* \geq W_F^R\right) \end{aligned} \quad (15)$$

Una parte de las diferencias salariales medias serían atribuibles a diferencias en las características socioeconómicas de hombres y mujeres, y otra a diferencias en los rendimientos de las características, este último factor estaría asociado a discriminación salarial.

Como se ha señalado anteriormente, la literatura de análisis de discriminación salarial se ha centrado sobre todo en analizar salarios observados y, por ello, ha existido y existe una amplia discusión en torno a la interpretación de la expresión (15), y en concreto, en torno a la descomposición primero, e interpretación después, de los términos ligados a la selección. A este respecto, el artículo de Neuman and Oaxaca (2004) resulta esclarecedor: en términos del análisis de la discriminación salarial, no existe una única forma de descomposición de esos términos, lo cual complica enormemente la interpretación de resultados. Al margen de esta cuestión, el marco conceptual presentado en este artículo deja claro que las diferencias salariales medias observadas, son fruto de la combinación no lineal de los determinantes de las ofertas salariales y de los determinantes de los salarios de reserva, con lo cual no es posible diferenciar fácilmente lo que es discriminación salarial pura por el lado de la oferta salarial, de lo es otro tipo de discriminación ligada a los salarios de reserva. De ahí el interés que puede tener en análisis de la discriminación salarial en el contexto de ecuaciones como (12) o (13) en dónde no existe ningún tipo de confusión al respecto.

En definitiva, en este trabajo sugerimos que se analicen las descomposiciones (12) a (15) y que de cada una de ellas se puede obtener información relevante para el análisis de la discriminación salarial por razones de género. Efectivamente, a partir de la expresión (12) y la correspondiente descomposición, se podría disponer de una medida global de discriminación salarial, libre de interferencias de cualquier posible elemento discriminatorio ajeno al proceso de oferta salarial, relacionado, en definitiva, con los salarios de reserva. La expresión (13) nos permitiría disponer de medidas de discriminación que se generan en la determinación del salario de reserva y que posiblemente tengan su origen en factores socioeconómicos, ajenos al menos en parte al mercado laboral, y que

determinan, por ejemplo, que una mujer casada con un niño menor de 1 año es posible que tenga un salario de reserva distinto que su marido. La expresión (14) nos permitiría disponer de medidas de discriminación conjunta salarios ofertados-salarios de reserva en el contexto de los determinantes de la observación salarial. En concreto, la distancia conjunta por género entre salario ofertado y salario de reserva, permitiría estudiar los determinantes de la observación o no de los salarios ofertados y hasta qué punto incide sobre dicha observación la discriminación. La expresión (15) permitiría disponer de medidas de discriminación salarial para el grupo de individuos que trabaja.

3. Aplicación para el caso de España

Con objeto de ilustrar las propuestas realizadas en el apartado anterior, se han analizado, empleando los datos para España de la muestra ampliada de PHOGUE 2000, las diferencias salariales entre hombres mujeres. La muestra está constituida por los mayores de 15 años y menores de 65 que no se dedican a actividades empresariales. El total de individuos considerados es de 11.364 hombres y 13.196 mujeres. Del total de individuos son asalariados 7.494 hombres y 4.684 mujeres y, para esta submuestra, se observa que los hombres tienen en media un salario hora que es un 14,5% superior al de las mujeres². En el Apéndice del artículo se presenta la descripción de las variables empleadas.

Con los 11.364 hombres y las 13.196 mujeres se han estimado, por separado para hombres y mujeres, modelos probit de determinación de la probabilidad de observar o no el salario que se presentan en las columnas (I) y (II) de la Tabla 1. Las variables incluidas en los

² Calculado como $\left(\overline{\log W_M} - \overline{\log W_F}\right) \times 100$.

modelos intentan recoger los factores económicos y sociales que pueden incidir sobre esa probabilidad. Por otro lado, con los 7.494 hombres y 4.684 mujeres asalariados, se han estimado las ecuaciones salariales mincerianas que se presentan en las columnas (III) y (IV) de la Tabla 1, que incluyen como regresores la educación, la experiencia y su cuadrado, el término de selección obtenido de los modelos probit y un conjunto de variables que indican la región de residencia de los individuos.

De los resultados presentados en la Tabla 1 hay que destacar que el término de selección resulta significativamente distinto de cero en las ecuaciones salariales, lo que se puede interpretar como evidencia a favor de la existencia de sesgo de selección y de que los salarios medios observados y ofertados pueden diferir tanto para hombres como para mujeres. De hecho, el signo negativo en el parámetro ligado al inverso del ratio de Mills en el caso de los hombres indica que el salario medio ofertado de un hombre que trabaja está por debajo del salario medio ofertado a un hombre elegido aleatoriamente. Para el caso de las mujeres, el signo positivo del parámetro ligado al término de selección indica que el salario medio ofertado para una mujer trabajadora está por encima del salario medio ofertado a una mujer elegida aleatoriamente.

A partir de los resultados recogidos en la Tabla 1 se han obtenido unos salarios medios ofertados y de reserva por género que se presentan en la Tabla 2. Esta tabla recoge también los salarios observados medios por género. Respecto a los resultados obtenidos, cabe señalar en primer lugar que, en consonancia con lo que se obtiene en Olivetti y Petrongolo (2008) y Blau y Kahn (2006), las diferencias salariales observadas por género son menores que las diferencias salariales en términos de salarios ofertados o potenciales, lo cual

significa que la selección juega un papel activo en el sentido de estrechar las diferencias salariales observadas respecto a las ofertadas. Nuestros resultados sugieren que ese estrechamiento tiene básicamente su origen en los efectos de la selección para el caso de las mujeres. Efectivamente, tal y como se puede apreciar, los salarios medios ofertados y observados de los hombres son muy parecidos. No ocurre lo mismo para las mujeres, donde hay diferencias en la media del 31%, que explican casi en su totalidad la diferencia por género entre salarios ofertados y observados.

En la Tabla 2 también se observa que los salarios medios de reserva de las mujeres son un 52,3% superiores a los de los hombres. Además, para los hombres, los salarios medios observados y ofertados están en media por encima de los salarios medios de reserva, mientras que, para las mujeres, ocurre lo contrario. Este resultado está en consonancia con el dispar comportamiento de las mujeres y los hombres en cuanto a la participación, que en nuestra muestra es del 65,9% para los hombres y del 35,5% para las mujeres.

A partir de los resultados de la Tabla 1 y empleando las expresiones (12) a (15) se pueden descomponer las diferencias salariales, lo que se recoge en la Tabla 3. En concreto, la columna (I) de dicha tabla recoge los resultados obtenidos para la descomposición de las diferencias en los salarios ofertados. La columna (II) de la Tabla 3 recoge dicha descomposición para los salarios de reserva y la columna (III) la descomposición para la diferencia salarios ofertados-salarios de reserva. Las columnas (IV) a (VII) recogen los resultados obtenidos para las diferentes descomposiciones de las diferencias salariales observadas, todas ellas presentadas en Neuman y Oaxaca (2004).

Cabe destacar de los resultados presentados en la Tabla 3 que independientemente de las diferencias salariales entre hombres y mujeres consideradas, el factor de discriminación tiene una mayor contribución relativa en la existencia de dichas diferencias que el factor de la diferencia de características. Este resultado está en consonancia con lo obtenido en los trabajos aplicados para España en los que con diferentes bases de datos y metodologías se obtiene que el componente de características tiene poco peso en la descomposición de las diferencias salariales (ver, por ejemplo, De la Rica y Ugidos, 1995, Hernández y Méndez, 2005, García, Hernández y López, 2001, De la Rica, Dolado y Llorens, 2008).

En términos de los salarios ofertados se obtiene que el componente de discriminación es de 0,396 (ver columna (I) Tabla 3). Este componente de discriminación en términos de salarios ofertados es complicado de comparar con el componente de discriminación para el caso de los salarios observados (columnas (IV)-(VII) Tabla 3). Como se ha comentado anteriormente, en el caso de salarios observados y problemas de selección no existe una única forma de realizar la descomposición de las diferencias salariales, lo cual complica enormemente la interpretación y comparación de resultados.

En la columna (II) de la Tabla 3 se observa que en términos de los salarios de reserva el componente de discriminación es muy importante (-0,568) y que las características de los individuos prácticamente no explican nada del diferencial en salarios de reserva. El signo negativo del componente de discriminación, dado que el salario de reserva medio de las mujeres es mayor que el de los hombres, puede interpretarse como evidencia de discriminación contra las mujeres en tanto en cuanto un mayor salario de reserva hace más difícil el acceso al mercado de trabajo. El origen de esa discriminación está posiblemente

en factores socioeconómicos que hacen que hombres y mujeres de iguales características tengan diferentes salarios de reserva.

En la columna (III) de la Tabla 3 se observa que en los determinantes de la observación o no del salario (comparación del salario ofertado con el de reserva) de nuevo el factor de discriminación es muy importante (0,964), lo que sugiere que en la diferente participación laboral de hombres y mujeres incide notablemente la discriminación tanto en los salarios ofertados como en los de reserva.

4. Conclusiones.

Los resultados de este trabajo sugieren que es fundamental distinguir entre los salarios ofertados y los observados a la hora de analizar las diferencias salariales por género. Las diferencias salariales observadas por género son menores que las diferencias salariales en términos de salarios ofertados o potenciales, lo cual significa que la selección juega un papel activo en el sentido de estrechar las diferencias salariales observadas. Nuestros resultados sugieren que ese estrechamiento tiene básicamente su origen en los efectos de la selección para el caso de las mujeres.

Respecto a la descomposición del diferencial salarial entre componente de características y de discriminación los resultados también muestran la importancia de realizar esta separación. En primer lugar porque en el análisis en términos de salarios ofertados no existe el problema de la descomposición del término de selección y, en segundo lugar porque el considerar tanto salarios observados como ofertados y de reserva aporta información sobre la discriminación salarial y su posible incidencia sobre la participación.

Apéndice. Definición de las variables

Las variables empleadas en el análisis se construyeron del siguiente modo:

Salario neto hora: Se construye a partir de la información disponible en PHOGUE sobre el número de horas trabajadas a la semana y los ingresos mensuales netos procedentes del trabajo por cuenta ajena. Se considera que el número de semanas que tiene un mes es 4,3452.

Educación: PHOGUE proporciona información sobre el nivel de estudios más alto completado por el individuo y asigna a cada nivel de estudios un valor numérico que recoge aproximadamente el número de años necesarios para completarlo. La variable empleada toma valor 2 para analfabetos y sin estudios, 5 para estudios primarios, 8 para primer nivel de secundaria, 9 para formación profesional de primer grado, 11 para la de segundo grado, 12 para segundo nivel de enseñanza secundaria, 15 para títulos universitarios de ciclo corto y 17 para título universitario de ciclo largo y postgrados.

Experiencia: Se construye, a partir de la información disponible en PHOGUE como la diferencia entre la edad del individuo y la edad en la que el individuo dice que comenzó su vida laboral. En la construcción de esta variable se impone que no supere la diferencia entre la edad de jubilación (65 años) y los años de estudio.

Otras características de los individuos:

Estado Civil: Es una variable que toma valor 1 para los individuos que están casados o mantienen una unión de hecho y 0 en caso contrario.

Región de residencia: PHOGUE agrupa en siete las posibles regiones de residencia: Noroeste (Galicia, Asturias y Cantabria), Noreste (País Vasco, Navarra, Rioja y Aragón), Madrid, Centro (Castilla y León, Castilla La Mancha y Extremadura), Este (Cataluña, Comunidad Valenciana y Baleares), Sur (Andalucía, Murcia, Ceuta y Melilla) y Canarias.

Para el análisis empírico se creó una variable ficticia para cada posible región de residencia que tomaba valor 1 si el individuo residía en dicha región y 0 en caso contrario.

Cuidado: En PHOGUE se les pregunta a los individuos si sus actividades diarias incluyen como quehacer no remunerado el cuidado de niños propios o ajenos o el cuidado de otros adultos. Sobre la base de esta información se creó una variable ficticia que tomaba valor 1 si el individuo cuida niños, adultos o ambos y 0 en caso contrario.

Número de hijos: PHOGUE no proporciona directamente información sobre el número de hijos de un individuo. Se ha construido a partir de la información proporcionada sobre las relaciones de parentesco entre los individuos del hogar, asignando a cada individuo como número de hijos todos aquellos individuos del hogar que figuraran como hijos carnales o adoptivos.

Ingreso neto del resto del hogar (en 1999): Se construye a partir de la información proporcionada por PHOGUE como diferencia entre los ingresos netos totales del hogar en 1999 y los ingresos netos totales del individuo en 1999.

Estar empleado: Se crea una variable ficticia que vale 1 para los asalariados y que toma valor 0 para los individuos que no trabajan (es decir cuyos ingresos mensuales del trabajo por cuenta ajena son cero), habiendo excluido de la muestra a los empresarios.

Referencias

- Altonji, J.G. y Blank, R. M. (1999), Race and Gender in the labor market, *Handbook of Labor Economics*, (Editores: O. Ashenfelter y D. Card), Elsevier Vol. 3, cap. 48, 3143-3249.
- Arrazola, M.J. y de Hevia, J (2008), “Three measures of returns to education: An Illustration for the case of Spain”, *Economics of Education Review*, 27, 266-275.
- Blau, F. y Kahn, L.M. (2006), “The US gender pay gap in the 1990s: Slowing convergence”, *Industrial an Labor Relations Review*, 60(1), 45-66.
- Blinder, A. S. (1973), “Wage Discrimination: Reduced Form and Structural Estimates”, *Journal of Human Resources*, 8, 436-55.
- De la Rica, S., Dolado, J.J. y Llorens, V. (2008), “Ceilings or floors?: Gender wage gaps by education in Spain”, *Journal of Population Economics*, 21, 751-776.
- De la Rica, S. y Ugidos, A. (1995), “¿Son las diferencias en capital humano determinantes de las diferencias salariales observadas entre hombres y mujeres?”, *Investigaciones Económicas*, vol. XIX (3), 395-414.
- Dolton, P.J. y Makepeace, G.H. (1986), “Sample selection and Male-Female earnings differentials in the graduate labor market”, *Oxford Economic Papers*, 38, 317-341.
- García, J., Hernández, P. J. y López, A. (2001), “How wide is the gap? An investigation of gender wage differences using quantile regression”, *Empirical Economics*, 26 (1), 149-168.
- Gronau, R. (1974), “Wage Comparisons-A Selectivity Bias”. *Journal of Political Economy*, 82(6), 1119-43.

- Hernández, P.J. y Méndez, I. (2005), “La corrección del sesgo de selección en los análisis de corte transversal de discriminación salarial por sexo: estudio comparativo en los países de la Unión Europea”, *Estadística Española*, vol. 47 (158), 179-214.
- Lewis, H.G. (1974), “Comment son Selectivity Biases in Wage Comparisons”, *Journal of Political Economy*, 82 (6), 1145-55.
- Mulligan C. B. y Rubinstein, Y. (2005), “Selection, investment, and women’s relative wages since 1975”, National Bureau of Economic Research, WP 11159.
- Neal, D. (2004), “The Measured Black-White Wage Gap among Women Is Too Small”, *Journal of Political Economy*, 112 (February), S1-S28.
- Neuman, S. y Oaxaca, R.L. (2004), “Wage decompositions with selectivity-corrected wage equations: A methodological note”, *Journal of Economic Inequality*, vol. 2(1), April, 3-10.
- Johnson, W R., Kitamura Y y Neal D. (2000), “Evaluating a simple method for estimating black-white gaps in median wages”, *American Economic Review Papers and Proceedings*, 90, 339-43.
- Oaxaca, R. (1973), “Male-Female Wage Differentials in Urban Labour Markets”, *International Economic Review*, 14, 693-709.
- Olivetti C. y Petrongolo B. (2008), “Unequal Pay or Unequal Employment? A cross – Country Analysis of Gender Gaps”, *Journal of Labor Economics*, 26(4), 621-654.
- Reimers, C.W., (1983), “Labor market discrimination against hispanic and black men”, *Review of Economics and Statistics*, 65(4), 125-141.

Tabla 1.

Modelo Probit de observación salarial (“Estar empleado”) y ecuación salarial

	Probit		Ecuación salarial	
	Hombres	Mujeres	Hombres	Mujeres
	(I)	(II)	(III)	(IV)
Constante	-4.348 (0.132)	-4.224 (0.132)	6.001 (0.033)	5.333 (0.062)
Educación	0.030 (0.004)	0.084 (0.003)	0.058 (0.001)	0.081 (0.003)
Edad	0.256 (0.007)	0.188 (0.007)	-	-
Edad²	-0.003 (0.00001)	-0.002 (0.00001)	-	-
Cuidado	-0.203 (0.048)	-0.126 (0.030)	-	-
Renta del resto del hogar	-5.48×10^{-8} (0.82×10^{-8})	-5.88×10^{-8} (0.73×10^{-8})	-	-
Número de hijos	-	-0.153 (0.020)	-	-
Renta del resto del hogar * Número de hijos	2.13×10^{-8} (0.58×10^{-9})	1.76×10^{-8} (0.43×10^{-8})	-	-
Estado civil	0.653 (0.042)	-0.251 (0.033)	-	-
Experiencia	-	-	0.020 (0.002)	0.032 (0.002)
Experiencia²	-	-	-0.0001 (0.00004)	-0.0005 (0.00005)
Término de selección	-	-	-0.184 (0.023)	0.143 (0.032)
χ^2_6 variables de región de residencia	96.572	142.058	186.179	83.489
SE	0.398	0.427	0.390	0.381
N	11,364	13,196	7,494	4,684

Notas: Cuidado es un variable ficticia que toma valor 1 si el individuo tiene a su cuidado niños y/o adultos. Entre paréntesis se presentan los Errores estándar, SE=Error Estándar de la regresión.

Tabla 2. Media del logaritmo de los salarios

	Hombres	Mujeres	Diferencia
Observados	6.838	6.693	0.145
Ofertados	6.856	6.383	0.473
De reserva	6.325	6.846	-0.523

Tabla 3. Descomposición de las diferencias salariales

	Oferta salarial		Salarios de reserva		Diferencia Oferta salarial-salarios de reserva		Salarios observados Neuman y Oaxaca (2004)							
	(I)		(II)		(III)		(IV)		(V)		(VI)		(VII)	
	Valor	%	Valor	%	Valor	%	Valor	%	Valor	%	Valor	%	Valor	%
Componente de discriminación	0.396	83.7	-0.568	108.6	0.964	96.8	0.327	225.5	0.404	278.6	0.404	278.6	0.160	110.3
Componente de características	0.077	16.3	0.045	8.6	0.032	3.2	-0.017	-11.7	-0.015	-10.3	-0.291	-200.7	-0.015	-10.3
Término de selección							-0.165	-113.8	-0.244	-168.3	0.032	22.1		
Diferencia salarial Total	0.473	100	-0.523	100	0.996	100	0.145	100	0.145	100	0.145	100	0.145	100

Nota: Los cálculos recogidos en las columnas (IV), (V), (VI) y (VII) se han realizado de acuerdo con las expresiones (11), (12), (13) y (14), respectivamente de Neuman y Oaxaca (2004). Los % están calculados sobre la diferencia salarial correspondiente.