



PREVALENCIA DE LA BRECHA DE GÉNERO EN LOS INGRESOS DEL PERÚ Y LIMA 2010 - 2017

Mijail Davis Huamán Romero

Universidad Torcuato Di Tella
Maestría en Economía, Departamento de Economía
Ciudad de Buenos Aires, Argentina
2018

Prevalencia de la Brecha de Género en los Ingresos del Perú y Lima 2010 - 2017

Tesis presentada como requisito para optar al título de:
Magíster en Economía

Autor:
Mijail D. Huamán Romero

Tutor:
Ph.D. Christiam M. Gonzales Chávez

Directora:
Ph.D. Marzia Raybaudi Massilia

Universidad Torcuato Di Tella
Maestría en Economía, Departamento de Economía
Ciudad de Buenos Aires, Argentina
2018

”Yo no deseo que las mujeres tengan poder sobre los hombres, sino sobre ellas mismas” .

Mary Wollstonecraft

Agradecimientos

A todos aquellos que siempre tienen la voluntad de contribuir en mi desarrollo profesional y sobre todo en mi desarrollo como ser humano.

Resumen

En esta investigación se utiliza una extensión de la descomposición de Oaxaca y Blinder para la diferencia del logaritmo de los ingresos entre hombres y mujeres, corregida por sesgo de selección, en el ámbito de Perú y Lima metropolitana con el propósito de identificar cuánto de la brecha de ingresos se debe a características observables como los años de educación o la experiencia laboral, y cuánto a variables no observables (*discriminación*). Los resultados muestran que la brecha de género en ingresos se ha ido reduciendo en el periodo del 2010 al 2017, tanto para Lima Metropolitana como para el Perú. Asimismo, para el Perú se encontró que la parte no explicada de la diferencia de ingresos entre hombres y mujeres es 97% del total de la diferencia (diferencia explicada más diferencia no explicada); y 67,6% para Lima Metropolitana.

Palabras clave: Brecha de genero, Brecha de salarios, Descomposición de Blinder - Oaxaca, Sesgo de selección.

Abstract

This paper use an extension of Oaxaca - Blinder decomposition for the difference in the logarithm of income between male and female, this extension is about a correction for selection bias for being employment or not, in the metropolitan area of Lima and Perú to determine how much of the wage gap is due to observable characteristics such as years of education or work experience, and how much to unobservable variables. It was found that the gender wage gap has been reduced from 2010 to 2017, for both areas, Lima and Peru. Likewise, the unexplained part of the decomposition is 97% of the total difference for Perú, and 67,6% for Lima.

Keywords: Gender Gap, Gender wage gap, Blinder - Oaxaca decomposition, Selection bias

Contenido

Agradecimientos	IV
Resumen	v
1. Introducción	2
2. Situación Actual de la Brecha de Género en Perú	3
2.1. Brechas de género: salarios, participación laboral y educación superior	4
3. Modelo	7
3.1. Sesgo de Selección	8
4. Los Datos	10
4.1. Sesgo de selección	10
4.2. Ecuaciones	10
5. Resultados de las Estimaciones	12
6. Conclusiones	19
A. Anexo: Estimaciones Adicionales	20
Bibliografía	22

1. Introducción

La teoría microeconómica clásica nos dice que el trabajador percibe un sueldo igual a su productividad marginal del trabajo, por lo que para la misma actividad que pueda realizar un hombre o una mujer, con las mismas o similares características y la misma productividad, la condición de género no debería afectar el salario de un grupo u otro. Es decir, la discriminación en salarios está presente cuando los pagos realizados por un empleador, dado un nivel determinado de características productivas, son sistemáticamente distintos para diferentes grupos demográficos ya sea por género, raza, condición de pobreza, etc.¹

En ese sentido, el objetivo de esta investigación es determinar cuánto de la diferencia en los salarios² entre hombres y mujeres en el Perú y en Lima Metropolitana se debe a diferencias observables (años de educación, experiencia laboral, trabajo informal, etc.) y cuánto por cuestiones no observables, siendo este último, considerado como una medida de ‘*discriminación*’. Dicho de otro modo, Ehrenberg y Smith (2016) definen la existencia de la discriminación cuando un trabajador, con características productivas similares a otro trabajador, es tratado de manera diferente por pertenecer a un determinado grupo³.

Por lo mencionado y con el fin de determinar cuánto de la diferencia de ingresos se debe a variables observables y cuánto a variables no observables, en ésta investigación se utiliza una variación del modelo de descomposición propuesto por Oaxaca (1973) y Blinder (1973), teniendo en cuenta el sesgo de selección existente en la muestra, donde los ingresos son observados únicamente para los trabajadores ocupados (variable truncada).

Cabe mencionar que los datos que se han usado para esta investigación provienen de la Encuesta Nacional de Hogares del Perú (ENAHO) elaborada por el Instituto Nacional de Estadística e Informática del Perú (INEI) para los años 2010-2017, esta encuesta contiene información sobre empleo, condiciones de vida de la población, salud, educación, etc.

¹Ehrenberg y Smith (2016)

²Específicamente los ingresos percibidos en la ocupación de la actividad principal del trabajador dependiente, según la definición del Instituto Nacional de Estadística e Informática del Perú.

³Los autores diferencian la existencia de la discriminación en salarios y la discriminación ocupacional

2. Situación Actual de la Brecha de Género en Perú

Existe una amplia literatura alrededor del problema de la desigualdad de género¹ y que es abordado desde varias disciplinas. En particular, las ciencias económicas suelen enfocarse en la desigualdad de ingresos, sus determinantes y las causas que podrían explicar la diferencia que existe entre ambos grupos (hombres y mujeres). En ese sentido, el Instituto Europeo de la Igualdad de Género (EIGE) ha identificado nueve brechas que tienen influencia en el desarrollo de la economía:

- Participación de las mujeres en el mercado laboral
- Brechas en educación superior
- Brecha de salarios entre hombres y mujeres
- Uso del tiempo
- Liderazgo en los negocios
- Liderazgo en la política
- Violencia contra la mujer
- Salud
- Migración

Asimismo, dentro de la literatura empírica se destacan los avances y aportes sobre la descomposición de Blinder (1973) y Oaxaca (1973). Dicha descomposición explica la diferencia de ingresos o salarios de dos grupos (hombres y mujeres, negros y blancos, etc).

Bajo el marco establecido, en las siguientes subsecciones se detallan algunas estadísticas relevantes sobre desigualdad de género en el Perú.

¹Rivera (2013) muestra un resumen interesante de las teorías de discriminación en las últimas décadas. Asimismo, Ehrenberg y Smith (2016), en el capítulo 12 de su libro, realizan un resumen más completo sobre género, raza y etnicidad en el mercado laboral. Por otra parte, el Instituto Europeo de la Igualdad de Género identifica las principales variables concernientes a la igualdad de género.

2.1. Brechas de género: salarios, participación laboral y educación superior

De las nueve brechas que el EIGE considera importantes, tres de estas son de crucial relevancia por su influencia directa con la economía:

1. Brecha de género en ingresos percibidos (*Ing*)
2. Brecha de género en la participación laboral (*PL*)
3. Brecha de género en educación superior en carreras de ciencias, tecnología, ingeniería y matemáticas (*CTIM*)

Gráfico 2-1.: Brecha de ingresos mensual en el Perú desde el 2004-2017 (porcentaje)



Fuente: Elaboración propia con información de la ENAHO - INEI

En el gráfico 2-1 se observa que los ingresos mensuales de los hombres en el Perú ascendió a 1,486 soles en promedio para el año 2017, mientras que en el caso de las mujeres fue de 1,186 soles. Si se estima la brecha de ingresos según la ecuación 2-1, se tiene que para el 2017 la brecha fue del 20,2%. Al mismo tiempo, este porcentaje es 9.7 puntos porcentuales (pp) menos que el registrado en el 2004 (29,9%). De este modo se ve que hubo una mejora en el periodo 2004 - 2017. Por otro lado, en el mismo gráfico, se puede ver que esta disminución en la brecha viene acompañado con el aumento del ingreso mensual tanto para hombres como mujeres. Aún en estas circunstancias, la brecha continúa con un valor considerable.

$$Brecha_{Ing} = \left(1 - \frac{Ing_{mujeres}}{Ing_{hombres}}\right) \times 100 \quad (2-1)$$

En el gráfico 2-2 se muestra la participación laboral para ambos grupos de estudio, donde se observa que la cantidad total de hombres y mujeres en la fuerza laboral ha ido aumentando en los últimos 14 años, siendo 9.4 millones de hombres y 7.4 millones el de mujeres en el 2017; y 7.4 millones de hombres y 5.6 millones de mujeres en el 2004. Si se estima la brecha de participación según la ecuación 2-2, en el 2017 fue del 21,2%. Sin embargo, la menor brecha registrada fue en el 2010 (20,9%). Como se puede apreciar en el periodo 2010-2017 hubo una reducción de la brecha, sin una tendencia clara, a pesar de ello siguen existiendo diferencias marcadas en participación laboral.

$$Brecha_{PL} = \left(1 - \frac{\#Trabajadores_{mujeres}}{\#Trabajadores_{hombres}}\right) \times 100 \quad (2-2)$$

Gráfico 2-2.: Brecha de participación laboral en el Perú desde el 2004-2017 (porcentaje)



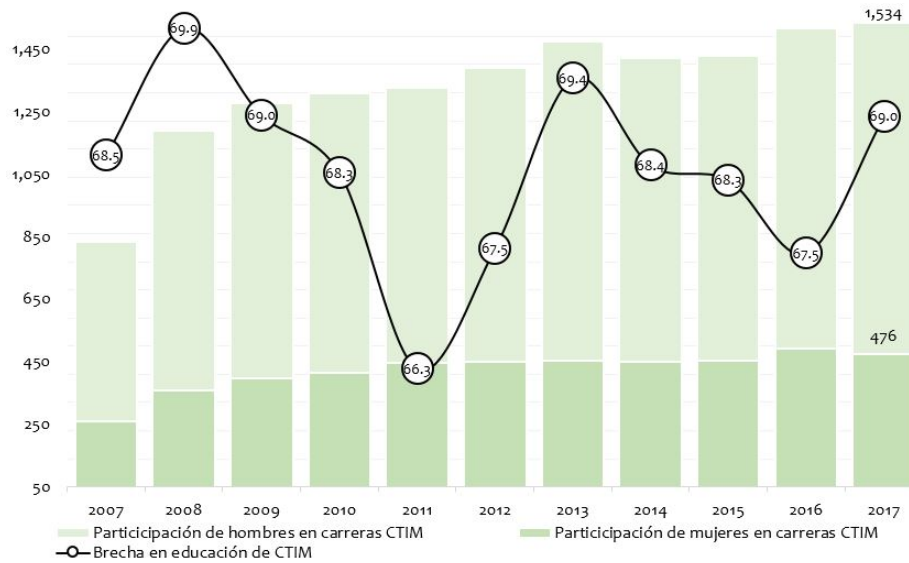
Fuente: Elaboración propia con información de la ENAHO - INEI

Relacionado con la brecha en educación superior, se estima la brecha de profesionales en ciencia, tecnología, ingeniería y matemáticas, entre hombres y mujeres, según la ecuación 2-3. La brecha de educación en carreras CTIM toma distintos valores entre 66.3% y 69.9% y no se aprecia una tendencia clara en el periodo 2007 - 2017. Sin embargo, la cantidad total de profesionales en CTIM ha ido aumentando, siendo la cantidad total de profesionales CTIM hombres 1.53 millones y 0.476 millones el de mujeres. Esta marcada diferencia de participación

entre hombres y mujeres en ramas de estudio que tengan que ver con las ciencias se debería a muchos estereotipos que limitan la participación de la mujer en tales carreras.

$$Brecha_{CTIM} = \left(1 - \frac{\#Profesionales_{mujeres,CTIM}}{\#Profesionales_{hombres,CTIM}} \right) \times 100 \quad (2-3)$$

Gráfico 2-3.: Brecha de educación superior en profesionales en ciencia, tecnología, ingeniería y matemática desde el 2007-2017



Fuente: Elaboración propia con información de la ENAHO - INEI

A modo de resumen, en el caso de la brecha de ingresos se observa una tendencia decreciente, mientras que en la brecha de participación laboral y educación, la tendencia no es tan clara.

3. Modelo

Se considera un modelo de regresión para el logaritmo de los ingresos mensuales ($\ln y$), específicamente, los ingresos mensuales de los trabajadores dependientes en la ocupación principal. Por lo que, se estima de manera separada para dos grupos $g = (h, m)$ la ecuación 3-1, del cual h representa al grupo de hombres y m al de mujeres:

$$\ln y_{ig} = X'_{ig}\beta_g + \varepsilon_{ig} \quad (3-1)$$

X'_{ig} contiene las variables explicativas de los ingresos, β_g el vector coeficientes asociados a cada variable explicativa y ε_{ig} el termino de error independientes e idénticamente distribuidas (iid) con media cero y varianza constante ($\sigma_{\varepsilon,g}$). Además, $i = 1, \dots, N_g$, y $\sum_g N_g = N$, siendo N_g la cantidad total de individuos del grupo g y N el total de individuos de la población. Para explicar la diferencia de ingresos entre hombres y mujeres se utiliza la descomposición propuesta por Blinder (1973) y Oaxaca (1973) que viene expresada por la ecuación 3-2 :

$$\overline{\ln y_h} - \overline{\ln y_m} = (\overline{X_h} - \overline{X_m})\hat{\beta}_h + \overline{X_m}(\hat{\beta}_h - \hat{\beta}_m), \quad (3-2)$$

donde los coeficientes de la ecuación 3-2 son estimados por mínimos cuadrados ordinarios por separado para cada grupo g . Sin embargo, Oaxaca y Ramsom (1994) proponen una descomposición general, donde se define un vector de coeficientes de no discriminación (β^*) como referencia al momento de estimar la diferencia de ingresos entre hombres y mujeres:

$$\overline{\ln y_h} - \overline{\ln y_m} = (\overline{X_h} - \overline{X_m})\beta^* + [\overline{X_h}(\beta_h - \beta^*) + \overline{X_m}(\beta^* - \beta_m)] \quad (3-3)$$

La primera parte de la ecuación $(\overline{X_h} - \overline{X_m})\beta^*$ explica la diferencia de ingresos en función de las variables explicativas u observables (X_g), mientras que la expresión entre corchetes de la ecuación es la parte no explicada y recoge la diferencia de ingreso entre hombres y mujeres explicada por variables no observadas. A este último componente la literatura suele atribuirle a la '*discriminación*'.

Por otra parte, existe una amplia discusión sobre el valor de β^* , por ejemplo, si $\beta^* = \beta_h$ se obtiene la expresión de la ecuación 3-2, es decir, el vector de coeficientes de no discriminación es el vector de coeficientes de la ecuación para lo hombres. Existen diversas alternativas y propuestas sobre el valor de los coeficientes de no discriminación, Reiner (1983) propone que $\beta^* = \beta^{0,5} = 0,5\beta_h + 0,5\beta_m$, mientras que Cotton (1988) sugiere la expresión de la ecuación 3-4:

$$\beta^* = \beta^{prop} = \frac{N_h}{N}\beta_h + \frac{N_m}{N}\beta_m \quad (3-4)$$

Adicionalmente, Neumark (1988) recomienda que se utilice el vector de coeficientes de una estimación apilada (*pooled regression*), que incluye a ambos grupos, como un estimado de $\beta^* = \beta^{pool}$. En este trabajo, se estima el valor de β^* para los tres casos mencionados, adicionalmente, se utiliza a modo de comparación, los valores de $\beta^* = \beta_h$ y $\beta^* = \beta_m$.

3.1. Sesgo de Selección

Una vez planteado el esquema general de la descomposición en el apartado anterior, se debe considerar la existencia de sesgo de selección presente en la población que percibe ingresos por trabajar al momento de estimar la brecha de ingresos entre hombres y mujeres, tal como lo plantean Neuman y Oaxaca (2004) ¹. Ello debido a que los ingresos de los trabajadores solo son observados para aquellas personas que participan del mercado laboral; por ello, esta es una variable truncada que requiere ser corregida. De este modo, se propone la función de empleo:

$$O_{ig}^s = Z'_{ig}\alpha_g + \mu_{ig}, \quad (3-5)$$

donde O_{ig}^s es la variable latente asociada a la condición de estar empleado y Z_{ig} es el vector de variables que determinan la condición de estar empleado. Ahora se asume que ε_{ig} y μ_{ig} son iid y tienen una función normal bivariada con media cero, varianza constante ($\sigma_{\varepsilon,g}$, $\sigma_{\mu,g}$) y covarianza ρ_g constante. Por lo tanto, la probabilidad de estar empleado viene expresado por la siguiente ecuación:

$$Prob(O_{ig}^s > 0) = Prob(\mu_{ig} > -Z'_{ig}\alpha_g) = \Phi(Z'_{ig}\alpha_g), \quad (3-6)$$

siendo $\Phi(\cdot)$ es la distribución acumulada de la distribución normal estándar. Por lo tanto, la esperanza del logaritmo de los ingresos para cada grupo g es determinada de acuerdo a la siguiente ecuación:

$$E(\ln y_{ig} | O_{ig}^s > 0) = X'_{ig}\beta_g + E(\varepsilon_{ig} | \mu_{ig} > -Z'_{ig}\alpha_g) = X'_{ig}\beta_g + \theta_g\lambda_{ig}, \quad (3-7)$$

donde $\theta_g = \rho_g\sigma_{\varepsilon,g}$, $\lambda_{ig} = \phi(Z'_{ig}\alpha_g)/\Phi(Z'_{ig}\alpha_g)$, y $\phi(\cdot)$ es la función de densidad normal estándar.

Finalmente, la ecuación a estimar incorporando el sesgo de selección es:

$$\ln y_{ig} | O_{ig}^s > 0 = X'_{ig}\beta_g + \theta_g\lambda_{ig} + error \quad (3-8)$$

¹Garavito (2011) estima un modelo de selección de Heckman en dos etapas para corregir el problema de sesgo de selección al momento de explicar la diferencia de ingresos por hora entre hombres y mujeres en el Perú para el año 2010

Para estimar la ecuación 3-8 se utiliza el procedimiento de Heckman en dos etapas, primero se estima con un modelo probit para la ecuación 3-6, posteriormente se calcula la inversa del ratio de Mills (λ_{ig}) para cada observación y se estima la ecuación 3-8. Como resultado, si se considera la existencia de sesgo de selección, la ecuación de descomposición de la brecha de ingresos entre hombres y mujeres 3-3 incorpora un nuevo término:

$$(\overline{\ln y_h} - \overline{\ln y_m}) = (\overline{X_h} - \overline{X_m})\beta^* + [\overline{X_h}(\beta_h - \beta^*) + \overline{X_m}(\beta^* - \beta_m)] + (\theta_h\lambda_h - \theta_m\lambda_m) \quad (3-9)$$

La última parte de la ecuación 3-9 ($\theta_h\lambda_h - \theta_m\lambda_m$) es la corrección por selección mencionada por Neuman y Oaxaca (2004).

4. Los Datos

Los datos utilizados en esta investigación fueron extraídos de la Encuesta Nacional de Hogares elaborado por el Instituto Nacional de Estadística e Informática del Perú para los años 2010-2017. En este caso, se consideran a los trabajadores dependientes cuyos ingresos mensuales provienen de la ocupación principal.

Se construyeron distintas variables como años de educación y experiencia laboral (variables de Mincer). Adicionalmente se crearon variables dicotómicas para el tipo de carrera o profesión de los individuos (Educación, Humanidades, ingeniería, etc), también se usó una variable que identifica si el centro de estudios fue o es privado o público, además se tuvo en consideración si la lengua materna del individuo fue español, extranjero o nativo (Quechua, Aymara, etc).

Garavito (2011) resalta la importancia del capital de las empresas en los salarios de los trabajadores, por consiguiente, se tomó una variable que identifica si el trabajador labora en un centro formal o informal, el tamaño de la empresa en función a la cantidad de trabajadores, además se crearon variables dicotómicas que identifica a que actividad económica se dedica la empresa del trabajador (Agropecuaria, Pesca, Minería, etc).

Por último se consideró una variable que identifica el área urbana o rural donde vive el trabajador.

4.1. Sesgo de selección

Para la ecuación de selección, las variables utilizadas como determinantes del empleo son edad, si tiene pareja o no, la incidencia de pobreza en el hogar que viven los trabajadores, si padece alguna enfermedad crónica, si el trabajador tiene hijos menores de 6 años y si tienen hijos entre 7 y 14 años. Asimismo, para la ecuación de selección se considera la subpoblación que pertenece a la Población Económicamente Activa (PEA), es decir, aquellas personas que están en capacidad de trabajar, y trabajan o desean hacerlo.

4.2. Ecuaciones

A modo de resumen, la ecuación 4-1 es la regresión para el logaritmo de los ingresos incorporando el parámetro asociado al sesgo de selección ($Mills^{-1}$), además del resto de variables explicativas. Del mismo modo, la ecuación 4-2 muestra la ecuación de selección con las varia-

bles explicativas que determinan la condición de estar ocupado o no, asimismo, cada ecuación se estima de manera separada para cada grupo de interés.

$$\begin{aligned}
 \ln y_{ig} = & \beta_{0,ig} + \beta_{1,ig} \text{años_educación} + \beta_{2,ig} \text{años_experiencia} + \beta_{3,ig} \text{años_experiencia}^2 \\
 & + \beta'_{4-12,ig} \text{carreras} + \beta_{13,ig} \text{centro_de_estudios} + \beta_{14-15,ig} \text{lengua_materna} + \\
 & \beta_{16,ig} \text{trabajo_informal} + \beta'_{17-23,ig} \text{actividad_económica} + \\
 & \beta'_{24-25,ig} \text{tamaño_empresa} + \beta_{26,ig} \text{urbano} + \theta_g \text{Mills}^{-1} + \text{error}
 \end{aligned} \tag{4-1}$$

$$\begin{aligned}
 \text{Prob}(O_{ig}^s > 0) = & \Phi(\alpha_{0,ig} + \alpha_{1,ig} \text{edad} + \alpha_{2,ig} \text{edad}^2 + \alpha_{3,ig} \text{pareja} + \alpha_{4,ig} \text{pobre} + \\
 & \alpha_{5,ig} \text{enfermedad_crónica} + \alpha_{6,ig} \text{niños_6} + \alpha_{7,ig} \text{niños_7_14})
 \end{aligned} \tag{4-2}$$

5. Resultados de las Estimaciones

Antes de presentar los resultados de las estimaciones de acuerdo al modelo planteado en esta investigación, algunos hallazgos interesantes a resaltar son los de Garavito (2011), donde encuentra que los varones en Perú obtienen un ingreso por hora 25 % mayor al de las mujeres para el 2010, sin embargo no distingue cuánto de esta diferencia se debe a diferencias observables y no observables. Otro resultado interesante es el de Atal, Ñopo y Winder (2010) en el caso de América Latina (incluido el Perú), donde la brecha de ingresos por género se encuentra entre 9 % y 27 %. Asimismo, los autores encuentran que luego de controlar por características observables, la parte no explicada de la brecha de ingresos por género alcanza el 20 % del salario promedio de las mujeres en la región. Con respecto al caso ecuatoriano, Rivera (2013) encuentra que la brecha de ingresos entre hombres y mujeres se han ido reduciendo de 15.1 % en el 2007 a 10.1 % en el 2012.

En cuanto a esta investigación, primero se realizan diversas estimaciones OLS para toda la muestra (hombres y mujeres), incorporando la variable género. Además se realiza una estimación por sesgo de selección en dos etapas (Heckman) tal como se muestra en la tabla **5-1**. La primera estimación, columna (1), de regresión lineal considera variables de Mincer, es decir, variables que describan las capacidades de los individuos para efectivamente trabajar, estos son los años de educación y de experiencia. El segundo modelo, columna (2), incluye variables que describen las características de la educación como son la carrera que ha estudiado y el tipo de centro de estudios donde ha estudiado, si es privado o público, y la lengua materna del trabajador. En el tercer modelo, columna (3), se incluyen las variables del primer modelo más variables relacionadas al centro de trabajo, como la actividad económica a la cual se dedica la empresa y su tamaño, es decir si es una mype, mediana o gran empresa, también se considera el área (urbano o rural) donde vive el trabajador. El cuarto modelo en la tabla **5-1**, columna (4), es la estimación que considera todas las variables anteriormente mencionadas. Por último, columna (5), se estima un modelo con ajuste por sesgo de selección, mediante el modelo de Heckman usando un modelo probit que muestra la probabilidad de estar ocupado en función de la edad, la edad al cuadrado, si posee pareja o no, si es pobre, si tiene enfermedades crónicas, si tiene niños menores de 6 años y si tiene niños menores de 14 años, pero mayores de 6.

De acuerdo a las estimaciones de la tabla **5-1**, se puede observar que la estimación donde se incorpora todas las variables, columna (4), posee un R cuadrado ajustado de 0.483, mayor al resto de estimaciones OLS. Asimismo, se observa que las mujeres ganan 24,41 % menos que los hombres de acuerdo a este modelo. Adicionalmente, las variables años de educación y experiencia tienen signo positivo y con efectos cuadráticos negativos para los años de expe-

Tabla 5-1.: Regresiones OLS (1) - (4), Heckman (5)

$\ln y$ (variable dependiente)	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) Heckman
Género (1 = mujer)	-0.3113***	-0.3298***	-0.2388***	-0.2441***	-0.2478***
Años educación	0.0794***	0.0632***	0.0337***	0.0287***	0.0340***
Años experiencia	0.0537***	0.0543***	0.0293***	0.0307***	0.0143**
(Años experiencia) ²	-0.0011***	-0.0011***	-0.0006***	-0.0006***	-0.0002
Carreras					
educación		0.0945***		-0.0500**	-0.0959
humanidades y arte		0.0871		0.0693	0.0843
CSCD ^{1/}		0.2309***		0.1169***	0.1294*
CNEC ^{2/}		0.0719*		-0.0557*	-0.0695
ingeniería, industria y construcción		0.2201***		0.0848***	0.0983
agropecuaria y veterinaria		0.0405		0.0321	0.0093
ciencias de la salud		0.2119***		0.1215***	0.1010
servicios personales		0.0709		0.0965	0.0982
Centro de estudios privado		0.0456***		0.0340**	0.0551
Lengua Materna					
nativo (quechua, aymara, etc)		-0.1198***		-0.0247	-0.0506
extranjeros (portugués, ingles, etc)		-0.0268		0.0292	0.0621
Trabajo informal			-0.4301***	-0.4244***	-0.3607***
Actividad económica de la empresa					
pesca			0.4382***	0.4370***	0.3785*
minería e hidrocarburos			0.7202***	0.7213***	0.6660***
manufactura			0.2222***	0.2203***	0.2221***
suministro de servicios básicos			0.2016***	0.1804***	0.1677
construcción			0.5084***	0.5150***	0.4503***
comercio			0.1679***	0.1544***	0.1880***
otros servicios			0.1765***	0.1780***	0.1683***
Tamaño de la empresa					
mediana empresa			0.3568***	0.3510***	0.3214***
gran empresa			0.4452***	0.4450***	0.4040***
Urbano			0.2818***	0.2666***	0.1909***
Constante	5.7601***	5.8977***	6.0240***	6.0613***	6.4274***
Ocupado					
Edad					0.0871***
Edad ²					-0.0009***
Pareja					0.1340***
Pobre					-0.1558***
Enfermedad crónica					-0.1912***
Niños 6					0.0343
Niños 7-14					0.0808***
Constante					-0.2904***
Mills					
Lambda					-2.4005***
Observaciones	26091	26088	25680	25677	27715
R^2 ajustado	0.295	0.306	0.478	0.483	

* $p < 0,05$, ** $p < 0,01$, *** $p < 0,001$

1/CSCD: Ciencias Sociales, Comerciales y Derecho

2/CNEC: Ciencias Naturales, Exactas y de la Computación

riencia. Por otra parte, incorporando la corrección por sesgo de selección, columna (5) de la tabla 5-1, se observa que las mujeres ganan 24,78 % menos que los hombres.

Sin embargo, las estimaciones de la tabla 5-1, reflejan que la condición de ser mujer implica 24,4 % menos de ingresos. Por ese motivo, se utiliza la descomposición de Oaxaca y Blinder para explicar cuánto de la diferencia de ingresos entre hombres y mujeres es por cuestiones observables como años de educación, experiencia, etc., y cuánto por cuestiones no observables.

Tabla 5-2.: Descomposición Oaxaca Blinder 2017, para diferentes valores de β^*

$\overline{\ln y} \setminus \beta^*$	(1) β^{pool}	(2) $\beta^{0,5}$	(3) β_h	(4) β_m	(5) β^{prop}
$\overline{\ln y_h}$	6.99537*** (0.000)	6.99537*** (0.000)	6.99537*** (0.000)	6.99537*** (0.000)	6.99537*** (0.000)
$\overline{\ln y_m}$	6.74346*** (0.000)	6.74346*** (0.000)	6.74346*** (0.000)	6.74346*** (0.000)	6.74346*** (0.000)
diferencia	0.25190*** (0.000)	0.25190*** (0.000)	0.25190*** (0.000)	0.25190*** (0.000)	0.25190*** (0.000)
descomposición					
diferencia explicada	0.02997** (0.003)	0.02813* (0.015)	0.02441* (0.015)	0.03186 (0.057)	0.02740* (0.012)
diferencia no explicada	0.22193*** (0.000)	0.22377*** (0.000)	0.22750*** (0.000)	0.22004*** (0.000)	0.22451*** (0.000)
observaciones	25677	25677	25677	25677	25677

p-values en paréntesis

* $p < 0,05$, ** $p < 0,01$, *** $p < 0,001$

La tabla 5-2 muestra las estimaciones que se realizaron tomando en cuenta la descomposición de Oaxaca-Blinder de la ecuación 3-3, y las variables explicativas de la ecuación 4-1 sin ningún ajuste por sesgo de selección (sin la variable $Mills^{-1}$). Como se puede apreciar, los resultados se muestran a nivel logarítmico y se hace una diferenciación entre la descomposición tanto la explicada como la no explicada. La descomposición explicada comprende las diferencias que son explicadas por las variables del modelo, como son los años de educación o de experiencia, etc. La descomposición no explicada no tiene un origen conocido y se entiende como aquellos aspectos que no están explicados por el modelo y que se puede etiquetar como ‘discriminación’. El primer modelo de la tabla, columna (1), estima la descomposición utilizando el vector de coeficientes de no discriminación de una regresión apilada (β^{pool}). El segundo modelo, columna (2), estima la descomposición con un vector de coeficientes de no discriminación ponderada de

igual manera para cada grupo g , es decir con un peso de 50 % para cada β_g de cada grupo g . El tercer y cuarto modelo utiliza como vector de no discriminación los coeficientes del grupo de hombres y mujeres, respectivamente. Por último, en la columna (5) se estima la descomposición tomando en cuenta como vector de coeficientes de no discriminación el β^* de la ecuación 3-4. Asimismo, se observa que la diferencia de ingresos ($\ln y$) entre hombres y mujeres es de 0.2519 y de esta diferencia la parte explicada oscila, para los diferentes valores de β^* , entre $0.024(\beta_h)$ y $0.031(\beta_m)$. Para el caso de la diferencia no explicada se puede ver que este valor esta alrededor de 0.22 el cual representa aproximadamente el 88 % de la diferencia total (0.2519).

Tabla 5-3.: Descomposición Oaxaca - Blinder 2017, corregido por sesgo de selección, para diferentes valores de β^*

$\overline{\ln y} \setminus \beta^*$	(1) β^{pool}	(2) $\beta^{0,5}$	(3) β_h	(4) β_m	(5) β^{prop}
$\overline{\ln y_h}$	6.99537*** (0.000)	6.99537*** (0.000)	6.99537*** (0.000)	6.99537*** (0.000)	6.99537*** (0.000)
$\overline{\ln y_m}$	6.74346*** (0.000)	6.74346*** (0.000)	6.74346*** (0.000)	6.74346*** (0.000)	6.74346*** (0.000)
diferencia	0.25190*** (0.000)	0.25190*** (0.000)	0.25190*** (0.000)	0.25190*** (0.000)	0.25190*** (0.000)
diferencia ajustada	0.34532*** (0.000)	0.34532*** (0.000)	0.34532*** (0.000)	0.34532*** (0.000)	0.34532*** (0.000)
descomposición					
diferencia explicada	0.01267 (0.191)	0.01324 (0.231)	0.00211 (0.821)	0.02438 (0.137)	0.01105 (0.284)
diferencia no explicada	0.33265*** (0.000)	0.33207*** (0.000)	0.34321*** (0.000)	0.32094*** (0.000)	0.33427*** (0.000)
observaciones	25677	25677	25677	25677	25677

p-values en paréntesis

* $p < 0,05$, ** $p < 0,01$, *** $p < 0,001$

La Tabla 5-3 muestra las estimaciones de la Descomposición de Oaxaca-Blinder para diferentes valores de β^* y considerando el ajuste por sesgo de selección según la ecuación 3-9 y las estimaciones de las ecuaciones 4-1 y 4-2. Se observa que, la diferencia ajustada, tomando en cuenta el sesgo de selección, de los ingresos entre hombres y mujeres es de 0.3452, mayor al 0.2519 de la estimación sin considerar el sesgo de selección. De este modo, la diferencia explicada de la descomposición, para los distintos valores de β^* , toma valores entre 0.00211

(β_h) y 0.0243 (β_m) siendo no significativos con p -values por encima del 0.137; mientras que la diferencia no explicada oscila entre 0.3432 (β_h) y 0.32094 (β_m) siendo significativos para todas las estimaciones con distintos β^* . Ahora bien, la diferencia ajustada no explicada de la brecha de ingresos entre hombres y mujeres es 97% del total de la diferencia ajustada, tomando como referencia $\beta^* = \beta^{prop}$.

En el cuadro 5-3, tomando como referencia el vector de coeficientes de no discriminación β^{pool} , $\beta^{0,5}$ y β^{prop} , la diferencia no explicada de los ingresos entre hombres y mujeres está alrededor de 0,33 y significativo para todos los casos, mostrando algún grado de robustez en dicha estimación de la diferencia no explicada.

Tabla 5-4.: Descomposición Oaxaca - Blinder 2010 - 2017, corregido por sesgo de selección, para $\beta^* = \beta^{prop}$

$\overline{\ln y}$ \ año	(1) 2010	(2) 2011	(3) 2012	(4) 2013	(5) 2014	(6) 2015	(7) 2016	(8) 2017
$\overline{\ln y_h}$	6.547*** (0.000)	6.661*** (0.000)	6.745*** (0.000)	6.808*** (0.000)	6.872*** (0.000)	6.916*** (0.000)	6.969*** (0.000)	6.995*** (0.000)
$\overline{\ln y_m}$	6.213*** (0.000)	6.361*** (0.000)	6.455*** (0.000)	6.534*** (0.000)	6.618*** (0.000)	6.670*** (0.000)	6.717*** (0.000)	6.743*** (0.000)
diferencia	0.334*** (0.000)	0.299*** (0.000)	0.290*** (0.000)	0.274*** (0.000)	0.254*** (0.000)	0.247*** (0.000)	0.252*** (0.000)	0.252*** (0.000)
diferencia ajustada	0.513*** (0.000)	0.443*** (0.000)	0.443*** (0.000)	0.422*** (0.000)	0.365*** (0.000)	0.332*** (0.000)	0.325*** (0.000)	0.345*** (0.000)
descomposición diferencia explicada	0.021 (0.195)	0.028* (0.032)	0.036** (0.003)	0.039*** (0.001)	0.012 (0.260)	0.007 (0.504)	-0.003 (0.754)	0.011 (0.284)
diferencia no explicada	0.492*** (0.000)	0.415*** (0.000)	0.407*** (0.000)	0.383*** (0.000)	0.353*** (0.000)	0.325*** (0.000)	0.328*** (0.000)	0.334*** (0.000)
observaciones	16048	18165	19804	24375	24322	24364	27339	25677

p -values en paréntesis

* $p < 0,05$, ** $p < 0,01$, *** $p < 0,001$

Por otro lado la tabla 5-4 muestra la evolución de la descomposición de la brecha de ingresos, desde el año 2010 hasta el año 2017, corregido por sesgo de selección para $\beta^* = \beta^{prop}$. La diferencia de los ingresos pasó de 0.334 en el 2010 a 0.252 en el 2017. En otras palabras, la diferencia de ingresos en soles, entre hombres y mujeres, como porcentaje del salario prome-

dio de las mujeres paso de 39,7% en el 2010 a 28,7% en el 2017 ¹. En cuanto a la diferencia ajustada por sesgo de selección, este pasó de 0.513 a 0.345 del 2010 al 2017. Vale la pena mencionar que la diferencia y la diferencia ajustada tienen tendencia al cierre de brechas. Con respecto a la diferencia explicada de la brecha de ingresos, solo es significativa para el 2011 (95%), 2012 (99%) y 2013 (0,99%), en el resto de años esta diferencia no es significativa. Mientras que la diferencia no explicada es significativa para todos los años y ha disminuido de 0.492 (2010) a 0.334 (2017)². Sin embargo, la diferencia no explicada como porcentaje de la diferencia ajustada se sigue manteniendo en valores altos, 96% para el 2010 y 97% para el 2017.

Tabla 5-5.: Descomposición Oaxaca - Blinder 2010 - 2017 para Lima Metropolitana y corregido por sesgo de selección, donde $\beta^* = \beta^{prop}$

$\overline{\ln y}$ \ año	(1) 2010	(2) 2011	(3) 2012	(4) 2013	(5) 2014	(6) 2015	(7) 2016	(8) 2017
$\overline{\ln y_h}$	6.845*** (0.000)	6.930*** (0.000)	7.008*** (0.000)	7.092*** (0.000)	7.170*** (0.000)	7.193*** (0.000)	7.255*** (0.000)	7.254*** (0.000)
$\overline{\ln y_m}$	6.447*** (0.000)	6.596*** (0.000)	6.706*** (0.000)	6.796*** (0.000)	6.882*** (0.000)	6.926*** (0.000)	6.988*** (0.000)	7.000*** (0.000)
diferencia	0.398*** (0.000)	0.334*** (0.000)	0.303*** (0.000)	0.297*** (0.000)	0.288*** (0.000)	0.267*** (0.000)	0.267*** (0.000)	0.254*** (0.000)
diferencia ajustada	0.413*** (0.000)	0.402*** (0.000)	0.416*** (0.000)	0.396*** (0.000)	0.318*** (0.000)	0.330*** (0.000)	0.344*** (0.000)	0.204*** (0.000)
Descomposición								
diferencia explicada	0.159*** (0.000)	0.126*** (0.000)	0.099*** (0.000)	0.122*** (0.000)	0.108*** (0.000)	0.092*** (0.000)	0.082*** (0.000)	0.066*** (0.000)
diferencia no explicada	0.254*** (0.000)	0.275*** (0.000)	0.316*** (0.000)	0.274*** (0.000)	0.210*** (0.000)	0.238*** (0.000)	0.263*** (0.000)	0.138*** (0.001)
observaciones	2927	3310	3449	4809	4879	4726	4886	4672

p-values en paréntesis

* $p < 0,05$, ** $p < 0,01$, *** $p < 0,001$

Del mismo modo, y con la finalidad de tener una muestra más homogénea, la tabla 5-5 muestra la descomposición de la diferencia de ingresos corregida por sesgo de selección para $\beta^* = \beta^{prop}$

¹Para calcular estos valores se aplica la función exponencial a la diferencia del ingreso entre hombres y mujeres, que en la tabla 5-4 están en logaritmos y se le resta uno.

²En el Anexo A-1 se muestra la descomposición de la diferencia de ingresos sin considerar la corrección por sesgo de selección.

en el ámbito de Lima Metropolitana desde el 2010 hasta el 2017. De modo similar a párrafos precedentes, la diferencia y diferencia ajustada de los ingresos entre hombres y mujeres viene cayendo en el periodo de análisis. En el caso de la diferencia ajustada, corregida por sesgo de selección, pasó de 0.413 a 0.204 del 2010 al 2017³. Esto significa que para el 2010, la diferencia ajustada en soles como porcentaje del ingreso promedio de las mujeres fue del 51,1 % en el 2010 y 22,6 % para el 2017. Asimismo, la diferencia explicada y no explicada estimados en la tabla 5-5 es significativa para todos los años. En esa misma línea, en el 2010, la diferencia no explicada (0.254) como porcentaje de la diferencia ajustada (0.413) fue del 61,5 %; mientras que para el 2017 fue del 67,6 %. Finalmente, para el caso de Lima Metropolitana, la diferencia ajustada en el 2017 (0,204) está por debajo de la diferencia sin ajuste (0,254) y también la diferencia no explicada fue del 0,138 menor al registrado un año antes (0,263), por lo que pareciera existir evidencia de una reducción de la diferencia de ingresos entre hombres y mujeres explicado por variables no observables ('discriminación').

³En el Anexo A-2 se muestra la descomposición de la diferencia de ingresos sin considerar la corrección por sesgo de selección para el ámbito de Lima Metropolitana.

6. Conclusiones

En base a los resultados encontrados, se aprecia que las estimaciones sin considerar el sesgo de selección subestiman la brecha de ingresos entre hombres y mujeres. Sólo para el caso de Lima Metropolitana en el año 2017, la diferencia de ingresos corregido por sesgo de selección (0.204) está por debajo de la estimación sin ajuste (0.254).

Asimismo, se halló que la brecha de ingresos en soles, como porcentaje del ingreso de la mujeres, cayó de 39,7 % en el 2010 a 28,7 % en el 2017, a nivel de Perú. Sin embargo, para el caso de Lima Metropolitana, este porcentaje pasó de 51,1 % a 22,6 % en el mismo periodo.

Finalmente, la diferencia no explicada de la brecha de salarios (*'discriminación'*) como porcentaje del total de la brecha, es muy elevado a nivel de Perú (97 % en el 2017); mientras que en el ámbito de Lima Metropolitana, este porcentaje es de 67,6 %.

A. Anexo: Estimaciones Adicionales

Tabla A-1.: Descomposición Oaxaca - Blinder 2010 - 2017, donde $\beta^* = \beta^{prop}$

$\overline{\ln y}$ \ año	(1) 2010	(2) 2011	(3) 2012	(4) 2013	(5) 2014	(6) 2015	(7) 2016	(8) 2017
$\overline{\ln y_h}$	6.547*** (0.000)	6.661*** (0.000)	6.745*** (0.000)	6.808*** (0.000)	6.872*** (0.000)	6.916*** (0.000)	6.969*** (0.000)	6.995*** (0.000)
$\overline{\ln y_m}$	6.213*** (0.000)	6.361*** (0.000)	6.455*** (0.000)	6.534*** (0.000)	6.618*** (0.000)	6.670*** (0.000)	6.718*** (0.000)	6.743*** (0.000)
diferencia	0.334*** (0.000)	0.299*** (0.000)	0.290*** (0.000)	0.274*** (0.000)	0.254*** (0.000)	0.247*** (0.000)	0.252*** (0.000)	0.252*** (0.000)
descomposición								
diferencia explicada	0.029 (0.090)	0.034* (0.015)	0.044*** (0.001)	0.053*** (0.000)	0.023* (0.044)	0.022* (0.045)	0.011 (0.292)	0.027* (0.012)
diferencia no explicada	0.306*** (0.000)	0.265*** (0.000)	0.246*** (0.000)	0.222*** (0.000)	0.231*** (0.000)	0.224*** (0.000)	0.240*** (0.000)	0.225*** (0.000)
observaciones	16048	18165	19804	24375	24325	24366	27344	25677

p-values en paréntesis

* $p < 0,05$, ** $p < 0,01$, *** $p < 0,001$

Tabla A-2.: Descomposición Oaxaca - Blinder 2010 - 2017 para Lima Metropolitana, donde $\beta^* = \beta^{prop}$

$\overline{\ln y}$ \ año	(1) 2010	(2) 2011	(3) 2012	(4) 2013	(5) 2014	(6) 2015	(7) 2016	(8) 2017
$\overline{\ln y}_h$	6.845*** (0.000)	6.930*** (0.000)	7.008*** (0.000)	7.092*** (0.000)	7.170*** (0.000)	7.193*** (0.000)	7.255*** (0.000)	7.254*** (0.000)
$\overline{\ln y}_m$	6.447*** (0.000)	6.596*** (0.000)	6.706*** (0.000)	6.796*** (0.000)	6.882*** (0.000)	6.926*** (0.000)	6.988*** (0.000)	7.000*** (0.000)
diferencia	0.398*** (0.000)	0.334*** (0.000)	0.303*** (0.000)	0.297*** (0.000)	0.287*** (0.000)	0.267*** (0.000)	0.267*** (0.000)	0.254*** (0.000)
descomposición								
diferencia explicada	0.168*** (0.000)	0.135*** (0.000)	0.108*** (0.000)	0.133*** (0.000)	0.121*** (0.000)	0.103*** (0.000)	0.095*** (0.000)	0.080*** (0.000)
diferencia no explicada	0.231*** (0.000)	0.199*** (0.000)	0.195*** (0.000)	0.164*** (0.000)	0.167*** (0.000)	0.165*** (0.000)	0.171*** (0.000)	0.174*** (0.000)
observaciones	2927	3310	3449	4809	4880	4726	4889	4672

p-values en paréntesis

* $p < 0,05$, ** $p < 0,01$, *** $p < 0,001$

Bibliografía

- [1] ATAL, J., ÑOPO, H. y WINDER, N. *New Century, old disparities: Gender and Ethnic gaps in Latin America*, Working paper series // Forschungsinstitut zur Zukunft der Arbeit, No. 5085, Institute for the Study of Labor (IZA), Bonn, 2010.
- [2] BEBLO, M., BENINGER, D., HEINZE, A. y LAISNEY, F. *Measuring Selectivity-Corrected Gender Wage Gaps in the EU*, ZEW, Mannheim, Université Louis Pasteur, Strasbourg, 2003.
- [3] COTTON, J. *On the Decomposition of wage Differentials*, Review of Economics and Statistics 70, pp 236 - 243, 1988.
- [4] EHRENBERG, R. y SMITH, R. *Gender, race, and ethnicity in the labor market*, Modern Labor Economics: Theory and Public Policy, twelfth edition, Chapter 12, pp 399 - 449, 2015.
- [5] EUROPEAN INSTITUTE FOR GENDER EQUALITY, «Economic Benefits of Gender Equality», *How the evidence was produced: briefing paper on the theoretical framework and model*.
- [6] ELBORGH-WOYTEK, K., NEWIAK, M., KOCHHAR, K., FABRIZIO, S., KPODAR, K., WINGENDER, P., CLEMENTS, B. y texscSchwartz, G., *Women, Work, and the Economy: Macroeconomic Gains from Gender Equity*, Strategy, Policy, and Review Department and Fiscal Affairs Department, International Monetary Fund, 2013.
- [7] GARAVITO, C. *Desigualdad en los Ingresos: Género y Lengua Materna*, Desigualdad Distributiva en el Perú, pag 236 - 265, 2011.
- [8] HECKMAN, J., *Sample Selection Bias as a Specification Error*, Econometrica, Econometric Society, vol. 47(1), pag 153 - 61, Enero 1979.
- [9] JANN, B. *A Stata implementation of the Blinder-Oaxaca decomposition*, ETH Zurich Sociology Working Paper Nro. 5, 2008.
- [10] NEUMAN, S AND R. L. OAXACA., *Wage decompositions with selectivity - corrected wage equations: A methodological note*, Journal of Economic Inequality 2: 3-10, 2004.
- [11] NEUMARK, *Employers discriminatory behavior and the estimation of wage discrimination*, Journal of Economic and Social Measurement, 24, pag. 55-61, 1988.

-
- [12] RAMANAYAKE, S.S. y GHOSH, T., *Role of Gender Gap in Economic Growth: Analysis on Developing Countries versus OECD Countries*, Indira Gandhi Institute of Development Research, Mumbai, 2017.
- [13] OAXACA, R., *Male-Female Wage Differential in Urban Labor Markets*, International Economic Review, Volume 14, Issue 3, 1973, 693-709.
- [14] OSPINO, C.G., ROLDÁN, P. y BARRAZA, N., *Oaxaca-Blinder Wage Decomposition: Methods, Critiques and Applications. A Literature Review*, Revista de Economía del Caribe Nro. 5, ISSN 2011-2106, 2009.
- [15] MACHADO, C., *Selection, Heterogeneity and the Gender Wage Gap*, Fundación Getulio Vargas, EPGE, 2011.
- [16] OTERO, J. V., *Descomposición Oaxaca-Blinder en Modelos Lineales y no Lineales*, Nro. 20, Instituto L.R. Klein - Centro Gauss, 2003.
- [17] REIMERS, C., *Labor Market Discrimination Against Hispanic and Black Men*, The Review of Economics and Statistics, Vol 65 Nro. 4, 1983.
- [18] RIVERA, J., *Teoría y Práctica de la Discriminación en el Mercado Laboral Ecuatoriano*, Analítika, Revista de análisis estadístico, Vol. 5(1): 7-22, 2013.
- [19] SLOCZYNSKI, T., *Average Gaps and Oaxaca Decompositions: A Cautionary Tale about Regression Estimates of Racial Differences in Labor Market Outcomes*, Brandeis University & IZA, 2018.