

PONTIFICIA UNIVERSIDAD CATÓLICA DEL PERÚ

FACULTAD DE CIENCIAS SOCIALES



Curso: Econometría 1

Integrantes: Ursula Cotrina Ccollana (20161708)
Daniela Estela Fernandez (20166140)
Sonia Asto Mercado (20166142)

Horario: 0722

Profesor: Luis García

Trabajo: Análisis de la ecuación de Mincer para el departamento de Cajamarca en
el año 2018

2020-2

Tabla de contenido

| | |
|---|-----------|
| 1. Introducción..... | 3 |
| 2. Marco Teórico Referencial sobre la ecuación de Mincer | 4 |
| 3. Análisis de datos | 5 |
| 4. Estimaciones Econométricas | 8 |
| 5. Conclusiones | 13 |
| 6. Referencias bibliográficas..... | 14 |

1. Introducción

A pesar del crecimiento económico continuo que ha experimentado el Perú en las últimas dos décadas, pues entre el 2007 y 2018, el ingreso promedio por trabajo creció en una tasa promedio anual del 5.7%, junto a esto, la fuerza laboral peruana se ha vuelto más experimentada y mejor educada, respecto a generaciones anteriores, esto no se ha reflejado en el cierre de la brecha salarial. Por ello, la brecha salarial sigue siendo un tema controversial en el país. Según el INEI (2018), la brecha salarial en el Perú asciende al 28,5%; es decir, el salario de las mujeres representó un 71.5% del salario de los hombres, por lo que la evidencia empírica refleja que los hombres perciben mayores salarios que las mujeres aun cuando no hay una gran diferencia en el grado de instrucción académica, Asimismo, de acuerdo al INEI (2018), la brecha salarial por género es mayor en el área rural (39.2%) que el área urbana (31.3%) al nivel nacional, entonces vemos que esta diferencia salarial de género mencionada anteriormente se acentúa más si evaluamos el lugar de residencia, específicamente si analizamos el área rural y el área urbana. Una manera de evidenciar estas cuestiones es a través de la ecuación de Mincer, la cual permite regresionar el salario (variable endógena) contra las variables exógenas correspondientes (educación, experiencia, sexo y región).

En el presente trabajo, planteamos la hipótesis de Mincer para el departamento de Cajamarca. Se evidenciará si la educación es una variable importante que determina los salarios; además, veremos si existe brecha salarial por género y lugar de residencia (urbano, rural). La estructura de la investigación será de la siguiente manera. En primer lugar, se explicará brevemente el marco teórico correspondiente a la ecuación de Mincer. En segundo lugar, se realizará un análisis de los datos recogidos desde el INEI, en el cual se describirán algunas características de la muestra utilizada para la investigación; asimismo, se explicarán las variables exógenas y la variable endógena que se utilizará para la estimación, de igual manera se realizará un análisis estadístico descriptivo de los datos. En tercer lugar, se presentarán los resultados obtenidos de las diferentes estimaciones econométricas realizadas; además, se mostrarán los resultados obtenidos de los test de identificación de heterocedasticidad (Test de White y Breusch-Pagan). Por último, se presentarán las conclusiones a las que se llega en la investigación.

2. Marco Teórico Referencial sobre la ecuación de Mincer

La ecuación de Mincer tiene sus inicios en el libro “Inversión en capital humano y distribución de la renta personal” en el año 1958. En él, se realiza un análisis económico que confirma que existe una relación positiva entre la educación y el salario. Los individuos más capacitados reciben un mayor salario y; por ende, experimentan un menor desempleo en comparación a los menos educados. Esta relación surge debido a la creencia tradicional de que la educación juega un rol importante en el estudio de la economía laboral y del capital humano. Por lo tanto, la función de ganancias de capital humano que desarrolla Jacob Mincer (1974) es la siguiente:

$$\ln(E_t) = \ln(E_0) + rs + \beta_1 t + \beta_2 t^2 + \mu$$

De la cual, el $\ln(E_t)$ es el logaritmo de los ingresos, $\ln(E_0)$ es la capacidad inicial de ganancia, s es el número de años de educación y su coeficiente r es la tasa de retorno de la educación. La experiencia laboral es la variable t , que es igual a $(A - s - b)$, donde A es la edad actual de la persona y b es la edad en la que inicia la escuela (comúnmente a los 6 años), por lo que $(s + b)$ sería la edad en que finaliza la escuela y, además, se espera que la persona trabaje inmediatamente luego de finalizar el colegio. Esta medición, según Mincer, se debe a que no existe una manera de observar directamente la capacitación adquirida en el trabajo (*on the job training*) o los años de experiencia laboral. Respecto a los coeficientes que acompañan tanto a la experiencia laboral como a la experiencia laboral al cuadrado, se espera que sean positivo y negativo, respectivamente. Esto se debe a que la experiencia al cuadrado indica la concavidad de los retornos del salario, o sea, que con el transcurrir del tiempo, los salarios aumentan, pero en menor proporción. El término de perturbación aleatoria es μ con distribución normal $(0, \sigma^2)$, que incluye a aquellas variables omitidas en la ecuación, las cuales afectan al salario del individuo.

La ecuación de Mincer utilizada es un modelo semilogarítmico y está descompuesta por la sumatoria de un término de educación lineal como exógena principal y variables de control como la experiencia, experiencia al cuadrado, variables dummies como sexo y región, y una dummy interactiva que relaciona el sexo con la educación. Se trabaja con la variable educación como una aproximación de la habilidad.

$$\ln(W_i) = \beta_1 + \beta_2 Educ_i + \beta_3 Exper_i + \beta_4 Exper_i^2 + \beta_5 Sexo_i + \beta_6 Sexo \times Educ_i + \beta_7 Region_i + u_i \dots \dots (1)$$

En cuanto a los parámetros, β_2 representa el incremento porcentual debido a un año más de educación, β_3 y β_4 representan el incremento porcentual debido a un año más de experiencia laboral, β_5 representa el incremento porcentual en el salario por ser hombre, lo cual manifestaría una discriminación en el mercado laboral, β_6 representa el impacto en el salario, debido a si se trata de hombre con educación vs a una mujer con educación. Finalmente, β_7 representa el incremento porcentual en el salario debido a si se reside en el área urbana.

3. Análisis de datos

La base de datos fue obtenida partiendo de los registros contenidos en la Encuesta Nacional de Hogares (ENAH) del 2018 elaborada por el Instituto Nacional de Estadística e Informática (INEI). Esta encuesta tiene como población objetivo a las viviendas particulares y a sus integrantes que se alojan a nivel nacional, incluyendo el área rural y área urbana. Asimismo, se extrajo los datos del módulo 02, 03 y 05 referentes a las características del miembro del hogar, educación y empleo e ingresos, respectivamente. Esto se realiza con la finalidad de obtener una base de datos de tipo corte transversa, la cual incluye una muestra de 99 646 personas a nivel nacional. Debido a que nos centraremos en realizar la estimación de los ingresos de la población del departamento de Cajamarca, el tamaño de la muestra se reduce a 4 326 personas. Asimismo, la población objetivo se limita a las personas que pertenezcan a la población económicamente activa (PEA), es decir, aquellas personas que tengan como edad 14 años a más y que sean asalariadas. Ante ello, la muestra final se reduce a 2012 observaciones, que viene a ser un 59,39% de la PEA (3 388).

De acuerdo con la ecuación (1), la variable “logaritmo natural del salario por hora”, $\ln(W_i)$, será la variable dependiente, la cual se obtiene de dividir los ingresos anuales entre el número de horas trabajadas multiplicado por las 48 semanas del año. Por otro lado, tenemos las variables independientes, las cuales están son “Años de educación” ($Educ_i$), “Años de experiencia” ($Exper_i$), “Años de experiencia al cuadrado” ($Exper_i^2$), “Sexo” ($Sexo_i$), “Sexo por educación” ($Sexo \times Educ_i$) y “Región” ($Region_i$).

Para determinar la variable “Años de educación” ($Educ_i$), se tomó en cuenta el nivel educativo previo aprobado de la persona, por lo que para aquellas personas que no tengan ningún nivel educativo aprobado, primaria incompleta o primaria completa, se les asignó 0 años de educación. En el caso de las personas que cuenten con secundaria incompleta o completa, se les asignó 6 años de educación; y para aquellas personas que tengan secundaria incompleta o completa, se les asignó 11 años de educación. Una vez obtenido los valores, se les sumó las variables “Último año que aprobó” o “Último grado que aprobó”, que corresponden a las preguntas 301B y 301C, respectivamente, del cuestionario del módulo 03. Por otra parte, la variable “Experiencia” ($Exper_i$) se calculó como la diferencia entre la edad de la persona y sus años de estudio. En cuanto a la variable “Sexo” ($Sexo_i$), esta se consideró como una variable dummy, la cual tomará el valor de 1 si es hombre y 0 si es mujer. De la misma forma, la variable “Región” tomará el valor de 1 si la persona pertenece al área urbana y 0 si pertenece al área rural. Finalmente, el término de perturbación se denota como u_i .

A continuación, se muestran los resultados obtenidos del análisis descriptivo. En Cajamarca, el 33.15% de la PEA entrevistada vive en el área urbana y el 66.85% vive en el área rural. Asimismo, el 53.36% de la PEA son mujeres y el 46.64% son hombres.

Gráfico 1

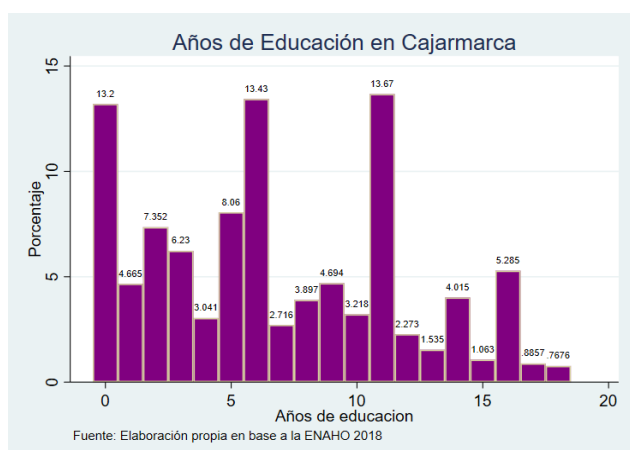
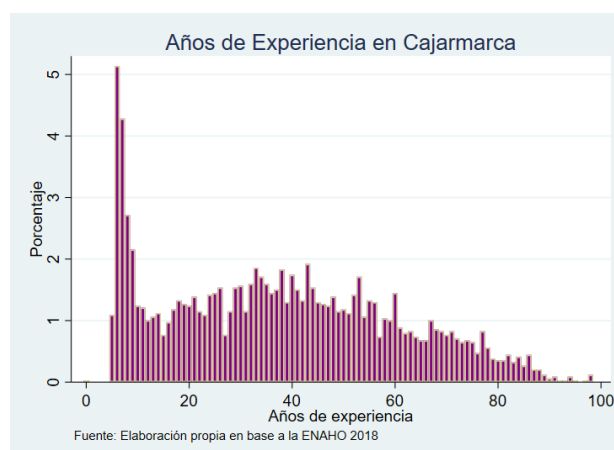


Gráfico 2



De acuerdo con el Gráfico 1, un mayor porcentaje de la población, 29.49%, finalizó la educación secundaria (11 años); de la misma forma, un 57.45% finalizó la educación primaria (6 años) que es el nivel promedio de años de estudio para toda la muestra. Por otra parte, la población que no tuvo años de formación educativa fue de un 13.2% y solo un 6.93% terminó sus estudios universitarios (16 años) y un 4.01% terminó sus estudios técnicos (14 años). En cuanto a los años de experiencia, en el Gráfico 2, se observa una mayor acumulación entre el rango de 20 y 60 años de experiencia, pues el 60% de la muestra tiene años de experiencia entre esos valores, por lo que los años de experiencia promedio es de 37 años; asimismo, un 34 % de la muestra tuvo entre 20 y 40 años de experiencia y un 27.8% tiene entre 0 y 20 años de experiencia.

Gráfico 3

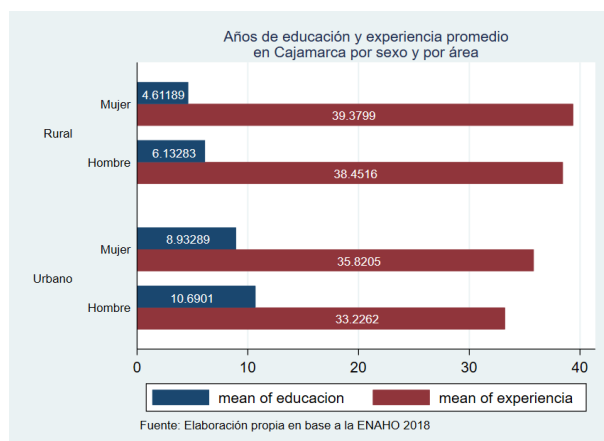
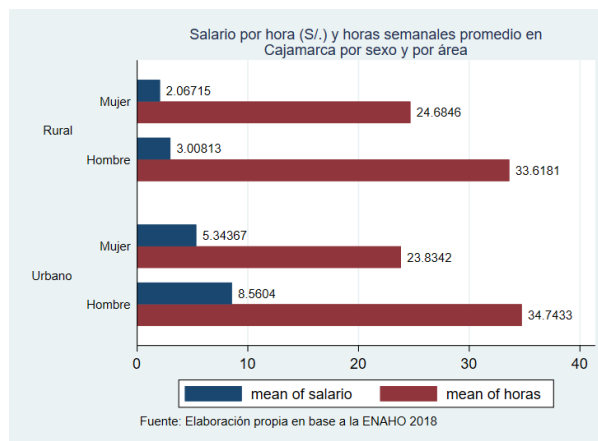


Gráfico 4



En el Gráfico 3, podemos observar que el promedio de años de educación es mayor en los hombres; quienes, en zona rural, tienen 1,52 años más de escolaridad en promedio que las mujeres y, en zona urbana, la diferencia es de 1,76 años de escolaridad. Sin embargo, las mujeres de zonas urbanas y rurales tienen más años de experiencia en promedio que los hombres, lo cual se debe a que tienen menos años de educación. También, se puede observar que el promedio de años de educación en mujeres de zona urbana es casi el doble que el de las mujeres que residen en zona rural, lo cual se debe al rol de la mujer en el

área rural. De la misma forma, para los hombres que viven en zonas urbanas, quienes a diferencia de los hombres que viven en zonas rurales, tienen 4,56 años más de educación en promedio; lo que nos indicaría que existe una brecha a nivel educativo por motivos demográficos. Por otra parte, en el Gráfico 4, se muestra una desigualdad salarial entre hombres y mujeres, tanto en las zonas urbanas como en las zonas rurales. Mientras que, los hombres de zonas urbanas ganan en promedio 8,56 soles por hora, las mujeres de zonas urbanas ganan 5,34 soles en promedio. De la misma forma, si multiplicamos el salario por hora con las horas semanales promedio obtenemos que la diferencia del salario semanal promedio entre hombres y mujeres que residen en la zona rural es de 50,1 soles; y el salario semanal promedio de los hombres que residen en áreas urbanas sería más del doble que el de las mujeres. Asimismo, una importante observación, es que las mujeres residentes de zonas urbanas ganan 76,34 soles más que las mujeres residentes de zonas rurales, en cuanto al salario semanal promedio. En conclusión, de acuerdo al Gráfico 4, la brecha salarial por género es mayor en el área urbana que en el área rural.

Gráfico 5

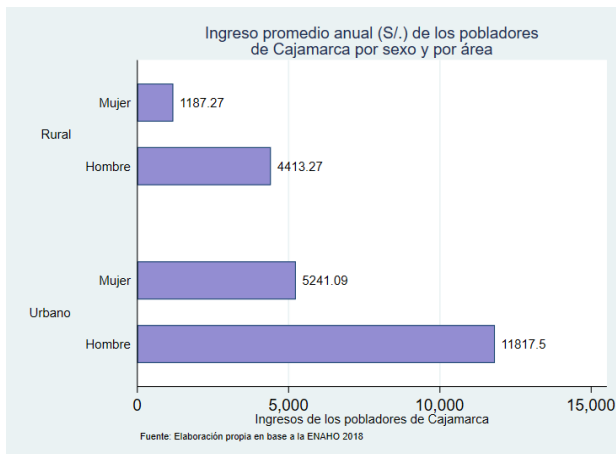
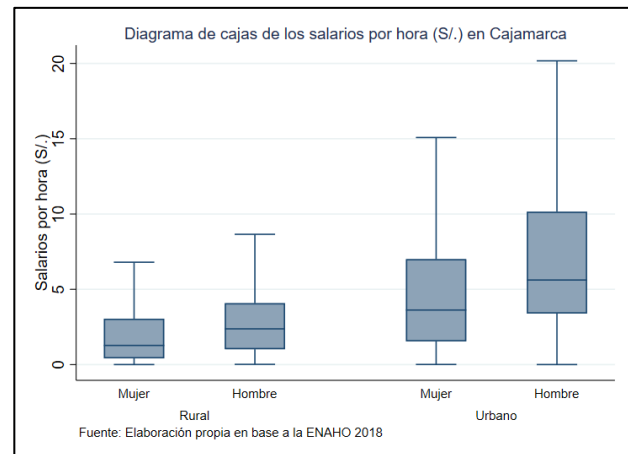


Gráfico 6



En el Gráfico 5, tenemos el ingreso promedio anual según sexo y región. Las mujeres tienen un ingreso promedio anual de S/. 2525.8, al comparar con el ingreso promedio anual de las mujeres por región, vemos que el promedio de las mujeres rurales (S/.1187.2) es menor, por lo que se encuentran en una situación más vulnerable. Del mismo modo, los hombres tienen un ingreso promedio anual de S/. 6878.2, y también se observa que los hombres rurales ganan por debajo del promedio (S/.4413.2). En ambos casos, vemos que en el área rural tiene un menor ingreso que el área urbana. Por otro lado, el ingreso anual promedio en el área rural es de S/2688.4, y en el área urbana es de S/8321.42, por lo que en ambos casos se observa que las mujeres reciben un menor salario que los hombres. En general, podemos decir que tanto en los casos presentados en el gráfico, los ingresos promedio mensuales no superan el sueldo mínimo (S/.930) a excepción de los hombres urbanos cuyo ingreso mensual promedio sería de (S/.984). En cuanto a los salarios por hora, en el Gráfico 6, se puede observar que existe una diferencia respecto a la dispersión de los salarios según región y sexo. En el caso de la mujer trabajadora, hay una mayor dispersión en el área urbana que rural, de igual manera para el caso de los hombres. Asimismo, también vemos que la mediana (línea dentro de la caja) del salario,

también varía según las categorías mencionadas, pues en el área rural tiene una menor mediana que el área urbana y de la misma manera, las mujeres tienen menor mediana que los hombres. Es importante recalcar, que se han omitido los outsiders (valores atípicos) para una mejor comprensión del gráfico.

4. Estimaciones Econométricas

La estimación se ha realizado por el método de Mínimos Cuadrados Ordinarios (MCO). Inicialmente, el modelo presentaba un problema de heterocedasticidad. Esto puede ser debido a que se trabaja con datos de corte transversal y no se podría suponer que las varianzas de los términos de perturbación sean iguales para todos los individuos de la muestra. La diversidad en los datos es muy variada, por ejemplo, puede depender del área de residencia, sexo, edad, etc, por lo que la variación de los salarios no tiene por qué ser uniforme, pues las personas con un mayor salario presentarán una mayor variabilidad. En concordancia a esto, en el Gráfico 6, se muestra que los hombres urbanos presentan una mayor dispersión y desviación estándar (ver Cuadro 2) que los otros grupos mencionados. Esto fue detectado mediante la prueba del test de White, mediante el cual se rechaza la hipótesis nula de homocedasticidad al 1% de significancia. Asimismo, mediante el test de Breusch-Pagan, la hipótesis nula de homocedasticidad es rechazada al 1% de significancia por casi todos los regresores, a excepción de la variable región, que rechaza dicha hipótesis al 5% de significancia como se observa en el cuadro 1. Este problema de heterocedasticidad ha sido solucionado mediante una estimación de MCO con errores robustos; es decir, usando el estimador consistente de la matriz de varianzas de White.

Cuadro 1

| Test de White | |
|-----------------------|----------|
| Variables | P-value |
| Todas | 0.0017 |
| Test de Breusch-Pagan | |
| Variables | P-value |
| Educación | 3.80E-08 |
| Experiencia | 8.70E-12 |
| Experiencia2 | 1.90E-11 |
| Sexo | 0.0016 |
| Sex_educ | 2.50E-07 |
| Región | 2.30E-05 |
| Todas | 8.50E-13 |

Elaboración Propia
Fuente: ENAHO 2018

Cuadro 2

| Región | Sexo | SD - Salario por hora |
|--------|--------|-----------------------|
| Rural | Mujer | 3.68 |
| | Hombre | 5.75 |
| Urbano | Mujer | 2.92 |
| | Hombre | 9.75 |

Elaboración Propia
Fuente: ENAHO 2018

En primer lugar, las regresiones de los distintos modelos presentados en la Tabla 1, fueron realizadas de manera separada. En cada modelo presentado, se va añadiendo una variable exógena respecto al modelo base de la ecuación de Mincer (Modelo 1), en el cual solo se regresa el salario en función de la educación, experiencia y experiencia 2, para así poder analizar si la inclusión de la variable independiente es adecuada o no.

Tabla 1

Estimación de la ecuación de ingresos de MINCER
Método: MCO con errores robustos

| Departamento: CAJARMARCA | Modelo 1 | Modelo 2 | Modelo 3 | Modelo 4 | Modelo 5 |
|-----------------------------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|
| Educación | 0.140*** (0.00654) | 0.135*** (0.00648) | 0.130*** (0.00878) | 0.0937*** (0.00956) | 0.0969*** (0.00756) |
| Experiencia | 0.000 0.0157*** (0.00513) | 0.000 0.0146*** (0.00511) | 0.000 0.0146*** (0.00510) | 0.000 0.0110** (0.00509) | 0.000 0.0111** (0.00509) |
| Experiencia2 | 0.002 -5.93e-05 (6.07e-05) | 0.004 -5.59e-05 (6.08e-05) | 0.004 -5.52e-05 (6.07e-05) | 0.030 -5.76e-05 (6.09e-05) | 0.030 -5.80e-05 (6.09e-05) |
| Sexo (hombre=1, mujer=0) | 0.329 | 0.359 0.330*** (0.0540) | 0.363 0.269*** (0.0927) | 0.344 0.381*** (0.0922) | 0.341 0.422*** (0.0541) |
| Sex_Educ | | 0.000 | 0.004 0.00904 (0.0102) | 0.000 0.00611 (0.00996) | 0.000 |
| Región (urbano=1, rural=0) | | | 0.374 | 0.540 0.623*** (0.0633) | 0.000 0.624*** (0.0633) |
| Constant | -0.731*** (0.128) | -0.853*** (0.130) | -0.823*** (0.135) | -0.695*** (0.133) | -0.715*** (0.128) |
| | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| Observations | 2,012 | 2,012 | 2,012 | 2,012 | 2,012 |
| R-squared | 0.204 | 0.219 | 0.219 | 0.256 | 0.256 |
| Adjusted R-Squared | 0.203 | 0.218 | 0.217 | 0.254 | 0.254 |
| RSS | 2765 | 2714 | 2713 | 2585 | 2585 |
| F-test | 188.1 | 148.7 | 122.2 | 130.3 | 153.1 |
| Prob>F | 0 | 0 | 0 | 0 | 0 |
| VIF | 10.99 | 8.510 | 8.260 | 7.220 | 7.220 |

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Sea ***, **, * los niveles de significancia al 1%, 5% y 10%

En la Tabla 1, vemos que se ha estimado 5 modelos, en los cuales se ha ido añadiendo variables progresivamente que pueden considerarse importantes al momento de la determinación del salario de una persona. Para ver si agregar dichas variables son beneficiosas o no para la estimación, analizaremos primero el R-cuadrado ajustado y luego, los Criterios de Información de Akaike y Schwarz. El cambio de los valores del R-cuadrado ajustado y de los Criterios de Información se observan en el Cuadro 3. En primer lugar, cuando incluimos las variables de sexo (modelo 2) y región (modelo 4), el R-cuadrado ajustado aumenta respecto al modelo base (modelo 1). En cambio, cuando el regresor sexo_educ (Modelo 3) es agregado, vemos que el R-cuadrado ajustado disminuye, por lo que no sería conveniente incluir este regresor, pero no es preocupante debido a que esta variable no es significativa. En segundo lugar, los criterios de Akaike y Schwarz refuerzan la conclusión ofrecida por el R-cuadrado ajustado, es decir, estos criterios disminuyen al agregar las variables de sexo y región, pero aumentan ligeramente en el caso de sexo_educ, pues el coeficiente obtenido no es significativo, pues tiene una significancia mayor a 10%. En conclusión, podemos decir que las variables de sexo y región son importantes por lo que es correcto que se hayan añadido al modelo porque ayudan a explicar mejor las variaciones del salario. Debido a que encontramos que la variable sexo_educ, no es significativa y según los Criterios sería una variable que no se debería incluir, se decidió analizar qué tanto cambiaría los coeficientes obtenidos del Modelo 4 si es que excluimos la variable sexo_educ. Respecto a ello, se ha obtenido que tanto el R-cuadrado ajustado como los Criterios de Información no cambian; sin embargo, los coeficientes de educación y sexo aumentan, otorgándole un mayor peso a estas como variables explicativas de la variación del salario.

Cuadro 3

| n | | k | SCR | R ajustado | Akaike (CIA) | Schwarz (CIS) |
|------|-----------------|---|------|---------------|--------------|------------------|
| 2012 | Modelo 1 | 3 | 2765 | 0.203 | 0.321 | 0.329 |
| | Modelo 2 | 4 | 2714 | 0.218 | 0.303 | 0.314 |
| | Modelo 3 | 5 | 2713 | 0.217 | 0.304 | 0.318 |
| | Modelo 4 | 6 | 2585 | 0.254 | 0.257 | 0.273 |
| | Modelo 5 | 6 | 2585 | 0.254 | 0.257 | 0.273 |

Elaboración propia

Fuente: ENAHO 2018

Como se observa en la Tabla 1, la variable endógena es el $\ln(\text{salario})$ y las variables con las que se regresa son la educación, experiencia, experiencia al cuadrado, sexo, sexo_educación y región. Los coeficientes estimados son significativos en su mayoría. Primero, en el caso de la variable educación, el parámetro es significativo al 1% en todos los modelos (del 1 al 5). Además, se evidencia que, en los modelos del 1 al 3, en promedio, la tasa de rentabilidad de la educación es de 13.5%, lo que indica que, por cada año adicional de educación, el salario por hora aumenta en un 13.5%; sin embargo, esto cambia para los modelos 4 y 5, ya que, por el incremento de un año de educación, el

salario aumenta aproximadamente solo en 9%. Segundo, en el caso de la variable experiencia, esta es significativa al 1% para los modelos del 1 al 3 y significativa al 5% para los modelos 4 y 5. En promedio, un aumento de un año de experiencia se ve reflejado en un incremento porcentual de 14% en el salario por hora para los modelos del 1 al 3 y un incremento de 11% para los modelos 4 y 5. Tercero, en el caso de la variable experiencia al cuadrado, su coeficiente es no significativo y es negativo para todos los modelos. El signo negativo es porque conforme se va obteniendo mayor experiencia, el impacto sobre el salario va disminuyendo, lo cual tiene sentido, debido a que conforme la edad va aumentando, las posibilidades de obtener un aumento en el salario disminuyen. Cuarto, si agregamos la variable sexo, su coeficiente es positivo y significativo al 1% en todos los modelos, lo cual significa que existe una diferencia de salarios entre mujeres y hombres a favor de estos últimos. En el modelo 5, aproximadamente, los hombres tienen un ingreso adicional de 42% sobre el ingreso de las mujeres, considerando la misma región. Quinto, al considerar la variable *sex_educ*, resulta que el parámetro es no significativo para el modelo 3 y 4, así que no ayuda a explicar las variaciones de salario. Por último, al agregar la variable región en el modelo 4, se observa que su coeficiente es positivo y significativo al 1%, por lo que existe un incremento en el salario del 62.3% solo por pertenecer al área urbana.

Se demuestra que existe una diferencia salarial entre hombres y mujeres en el modelo 4. Además, vemos que el coeficiente es positivo, por lo que los hombres ganan un salario mayor en 38% que las mujeres en Cajamarca. Esto va de acuerdo con lo esperado, pues según estudios sobre la brecha salarial en el Perú, los hombres reciben un salario mayor al de las mujeres. Según la INEI (2018), la brecha de ingresos por trabajo en Cajamarca ascendería al 37,8%; es decir, el salario de las mujeres representó un 62,2% del salario de los hombres. Asimismo, la brecha salarial estimada para Cajamarca (38%) se encuentra por encima de la brecha salarial para el Perú (28.5%), por lo que se puede concluir que Cajamarca está en una situación mayor de desigualdad salarial que el promedio.

Asimismo, respecto al R-cuadrado, su valor es muy similar al R-cuadrado ajustado y es menor a 0.30, como es de esperarse, ya que se trata de datos de corte transversal que incluye varios factores que influyen en la variación del salario. En el modelo 4, se puede afirmar que, en conjunto, las variables educación, experiencia, *experiencia2*, sexo, *sex_educ* y región explican un 25.4% las variaciones del salario. Por un lado, al analizar la significancia conjunta, cuando se realiza la prueba F con la hipótesis nula de que los coeficientes de los regresores sean cero, se obtiene que se debe rechazar dicha hipótesis, pues el $F - estadístico > F_{(0,95)}(q, n - k)$ para todos los casos. Por otro lado, respecto al factor de inflación de varianzas (VIF), hemos calculado la media del VIF de los modelos estimados. En todos los casos, a excepción del modelo 1, el valor del VIF es menor a 10, por lo que la multicolinealidad en esos modelos no es un problema, especialmente en el modelo 4 que tiene un VIF menor en comparación a los demás modelos. Se debe notar que las variables experiencia y experiencia al cuadrado elevan el valor de la media del VIF, pues presentan un valor VIF mayor a 15, mientras que el VIF

de las demás variables no superan el 4.5. Esto se debe a la relación existente entre dichas variables.

En la estimación econométrica se considera la educación como un aproximado de la habilidad, debido a que esta no se puede medir; sin embargo, al no ser una variable observable, puede estar incluida en el error (u_i) y ocasionar un problema de “endogeneidad” del regresor educación; es decir, que exista una correlación entre la variable educación y el error. La hipótesis de este problema sería que una mayor habilidad implicaría un mejor grado de instrucción académica. Si esto fuera verdad, entonces se podría concluir que los estimadores hallados no son consistentes, ya que habría una correlación entre un regresor y el error. Esto podría ser solucionado mediante la inclusión de una variable instrumental que no esté correlacionada con el error, pero sí con la educación. Ante ello, una forma de estimar parámetros consistentes es a través del método de Mínimos Cuadrados en Dos Etapas (MC2E). Las variables instrumentales que podríamos añadir son la educación del padre o madre, o incluso, el promedio de años de educación de la familia de un individuo que tenga 14 años o más, puesto que la educación de un individuo se encontraría relacionada con la educación de su familiar y esta, a su vez, no guardaría tanta relación con los salarios del individuo.

5. Conclusiones

En síntesis, podemos mencionar que se confirma la ecuación de Mincer planteada inicialmente, ya que la educación sí genera un impacto positivo en el sueldo. Las variables que no fueron significativas para explicar el modelo fueron sex_educ y experiencia2. Las variables de educación, sexo, región son significativas al 1% y la experiencia es significativa al 5%. Por otro lado, se corrobora que existe una brecha salarial, debido al sexo y lugar de residencia, siendo esto consistente con otros estudios que analizan la brecha salarial.

Se comprueba que el impacto de la experiencia en la determinación del salario es positivo, así como se confirma que la variable experiencia2 tiene un efecto negativo en las variaciones de salario. Esto debido a la concavidad de la variable. Asimismo, se evidencia que podría existir un problema de endogeneidad de los errores, debido a que la educación podría tener una correlación con la habilidad de la persona, y esta se encuentra en el error.

En el modelo 4 en el que se estima la ecuación (1), tanto el R-cuadrado y el R-cuadrado ajustado son 25,4% y 25,6%, respectivamente, por lo que la variación del salario no depende solo de las variables explicativas mencionadas sino también podría depender, por ejemplo, del tipo de trabajo (independiente o dependiente), del empleo (formal o informal), de la rama de actividad (comerciante, minero, agricultor, etc.), estado civil que se incluirían como una variable dummy en el modelo, entre otras variables. La inclusión de estas variables es justificable, por ejemplo, para el año 2018, la tasa de empleo informal fue de 89.3%. Por otro lado, la tasa de subempleo para el mismo año fue de 69.5% de la PEA (MTPE, 2019). En el caso para analizar la brecha salarial del empleo formal o informal se podría estimar mediante MCO e incluir la variable como dummy; sin embargo, hay otros métodos para tener en cuenta esta variable explicativa como Método de descomposición paramétrica de Oaxaca (1973).

Por último, se observó que las mujeres que residen en zonas rurales y urbanas de la región tienen mayores años de experiencia que los hombres; debido a que, tienen menores años de educación. Ello se puede evidenciar en los datos que brinda el INEI (2019) sobre el nivel de educación alcanzado en la región de Cajamarca en el 2018, en donde el 25,6% de la población de mujeres con edad de 15 años a más, alcanzan el nivel de secundaria; a diferencia de los hombres, donde el 38,2 % alcanza el mismo nivel. Para una población con edad de 25 años a más, solo el 17% de mujeres alcanzó el nivel de secundaria; mientras que el porcentaje de los varones fue de 28,2%. Por ello, es sustancial la inversión en el sector educativo con un enfoque en la calidad; ya que, de acuerdo con Mincer (1974), las capacidades para aprender en el centro de trabajo se encuentran relacionadas con las habilidades adquiridas en el colegio y una desigualdad en la inversión de capital humano significaría una desigualdad en los salarios.

6. Referencias bibliográficas

INEI. (2018a). *Encuesta nacional de hogares sobre condiciones de vida y pobreza*. Lima. Recuperado de <https://www.inei.gob.pe/bases-de-datos/>

INEI. (2018b). *Encuesta Nacional de Hogares: Diccionario*. Recuperado de <https://www.inei.gob.pe/bases-de-datos/>

INEI. (2019). *Perú: Evolución de los indicadores de empleo e ingreso por departamento 2007-2018*. Recuperado de

https://www.inei.gob.pe/media/MenuRecursivo/publicaciones_digitales/Est/Lib1678/libro.pdf

INEI. (2019). *Perú: Indicadores de Educación por Departamentos, 2008-2018*. Lima. Recuperado de https://www.inei.gob.pe/media/MenuRecursivo/publicaciones_digitales/Est/Lib1680/libro.pdf

Mincer, J. (1958). Investment in Human Capital and Personal Income Distribution. *Journal of Political Economy*, 66, 281–302.

Mincer, J. (1974). *Schooling, Experience, and Earnings*. New York: National Bureau of Economic Research. Recuperado de

<https://www.nber.org/books-and-chapters/schooling-experience-and-earnings>

Rios, P. (2019). *Educación y Brecha Salarial de Género en el Perú*.

Anexos

Ver Do file adjunto “Ecuación_Mincer_Cajamarca.do”