



UNIVERSIDAD CÉSAR VALLEJO

ESCUELA DE POSGRADO

**PROGRAMA ACADÉMICO DE MAESTRÍA EN
DOCENCIA UNIVERSITARIA**

**El Nivel Educativo y el uso de las Tecnologías de Información y
Comunicación y su impacto en los Ingresos de los egresados
universitarios en el Perú**

TESIS PARA OBTENER EL GRADO ACADÉMICO DE:

Maestro en Docencia Universitaria

AUTOR:

Canales Jerí, Juan Angel (ORCID: 0000-0002-8106-7500)

ASESORA:

Mgtr. Josco Mendoza, Janet Cenayra (ORCID: 0000-0001-8544-269X)

LÍNEA DE INVESTIGACIÓN:

Innovaciones Pedagógicas

LIMA – PERÚ

2022

Dedicatoria

A Ceci, Angello, Gino, y Luigui

por su amor y paciencia

Agradecimiento

A Jesús, a la vida, a la familia,
a mis amados libros
y ahora al internet y a la computadora.

Índice de Contenidos

Dedicatoria	ii
Agradecimiento	iii
Índice de Contenidos	iv
Índice de Tablas	v
Índice de Gráficos y Figuras	vii
Resumen	viii
Abstract	ix
I. INTRODUCCIÓN	1
II. MARCO TEÓRICO	8
III. METODOLOGÍA	21
3.1 Tipo y diseño de investigación	21
3.2 Variables y operacionalización	22
3.3 Población, muestra, muestreo, unidad de análisis	26
3.4 Técnicas e instrumentos de recolección de datos	27
3.5 Procedimiento	28
3.6 Método de análisis de datos	30
3.7 Aspectos éticos	32
IV. RESULTADOS	34
V. DISCUSIÓN	56
VI. CONCLUSIONES	62
VII. RECOMENDACIONES	63
REFERENCIAS	64
ANEXOS	68

Índice de Tablas

Tabla 1	Tenencia de infraestructura básica de las TICs de egresados universitarios	4
Tabla 2	Uso de los equipos y el internet de los Egresados Universitarios	4
Tabla 3	Variable dependiente Log (Ing), dimensión e indicador	23
Tabla 4	Variable independiente Educ, dimensión e indicador	23
Tabla 5	Variable independiente Edad, dimensión e indicador	24
Tabla 6	Variable independiente Género, dimensión e indicador	25
Tabla 7	Enaho, número de preguntas e ítems	26
Tabla 8	Síntesis de los temas investigados por el enaho	28
Tabla 9	Módulos del enaho a analizar	29
Tabla 10	Data del enaho a analizar por variables de la investigación	29
Tabla 11	Resultados del procesamiento de las variables de la investigación: Ing	30
Tabla 12	Resultados del procesamiento de las variables de la investigación: Educ	34
Tabla 13	Resultados del procesamiento de las variables de la investigación: Edad	35
Tabla 14	Resultados del procesamiento de las variables de la investigación: Sexo	35
Tabla 15	Creación de las variables de la investigación: Log (Ing)	36
Tabla 16	Creación de las variables de la investigación: Edad ²	37
Tabla 17	Tabla de contingencia: Sexo y Educ	37
Tabla 18	Estadísticas básicas de las variables del modelo	40
Tabla 19	Matriz de correlación entre variable dependiente y variables independientes	42
Tabla 20	Resultados de la regresión del modelo (2)	42
Tabla 21	Test de Breusch – Pagan	43
Tabla 22	Test de factor de inflación de la varianza (vif)	44
Tabla 23	Regresión robusta del modelo (2)	45
Tabla 24	Estadísticas básicas de las variables cuantitativas del modelo (2)	46
Tabla 25	Tabla cruzada de Educ y Sexo a nivel población	48

Tabla 26	Resultados de la regresión del modelo (2)	49
Tabla 27	Estadísticas básicas de las variables del modelo (3)	50
Tabla 28	Matriz de correlación entre variable dependiente y variables independientes	51
Tabla 29	Resultados de la regresión del modelo (3)	52
Tabla 30	Test de Breusch – pagan modelo (3)	53
Tabla 31	Test de factor de inflación de la varianza (vif) del modelo (3)	53
Tabla 32	Regresión robusta del modelo (3)	55
Tabla 33	Estadísticas básicas de las variables cuantitativas del modelo (2)	56
Tabla 34	Tabla cruzada de Educ y Sexo a nivel población	57
Tabla 35	Resultados de la regresión del modelo (2)	58
Tabla 36	Resultados de la tabla resumen de contingencia: Educ y TICs del modelo (3)	59
Tabla 37	Resultados de la regresión del modelo (3)	60
Tabla 38	Resultados comparativos de las regresiones de los modelos (2) y (3)	61

Índice de Gráficos y Figuras

Figura 1	Resultados del procesamiento de las variables de la investigación: Sexo	36
Figura 2	Pruebas gráficas de normalidad: log (Ing)	38
Figura 3	Pruebas gráficas de normalidad: Educ	38
Figura 4	Pruebas gráficas de normalidad: Edad	39
Figura 5	Pruebas gráficas de normalidad: Sexo	39
Figura 6	Gráfico de la matriz de dispersión	41
Figura 7	Gráfico de Kernel, normalidad de los residuos	45
Figura 8	Resultados del uso del internet	47
Figura 9	Resultados del procesamiento de las variables de la investigación: TICs	48
Figura 10	Gráfico de la matriz de la dispersión de las variables del modelo (3)	51
Figura 11	Gráfico de Kernel, normalidad de los residuos del modelo (3)	54

Resumen

Esta investigación tiene por objeto determinar el impacto, del nivel educativo logrado, la experiencia y el género de los egresados universitarios en el nivel de ingresos que perciben en su actividad laboral, para ello el marco conceptual que se ha utilizado es el de la teoría del capital humano, en especial la ecuación de Mincer para el trabajo de la evidencia empírica.

La metodología que se ha seguido es el de la ruta cuantitativa, no experimental, de tipo transeccional o transversal, específicamente de tipo correlacional o causal, en el que se ha aplicado la estimación de la regresión por mínimos cuadrados ordinarios en dos etapas la primera asumiendo como variable dependiente al logaritmo del ingreso y como variables independientes al nivel educativo logrado, a la experiencia y al tema de género y en una segunda adicionando al modelo propuesto la variable tics referida al uso o no de la computadora y si este hecho impacta en los ingresos.

Los resultados obtenidos han confirmado las hipótesis de la investigación, el nivel educativo impacta por cada año adicional de estudio positivamente en 11% en los ingresos, a más experiencia impacta positivamente en 7% en los ingresos, pero el hecho de ser mujer impacta de forma diferenciada, las mujeres perciben en menos 39.8% respecto que los hombres.

En la segunda etapa del modelo conceptual se concluye que si aumenta en un año la educación, el impacto en los ingresos es el aumento en 9%, a más experiencia esto impacta en los ingresos en 8%, por ser mujer los ingresos son menores en 41% respecto a lo que perciben los hombres y si utilizan TICs esto impacta positivamente en sus ingresos, en 40%, que si no lo utilizan.

Palabras clave: Ingresos, experiencia, sexo, tics, teoría del capital humano

Abstract

The purpose of this research is to determine the impact of the educational level attained, the experience and gender of university graduates on the level of income they receive in their labor activity, for which the conceptual framework used is that of human capital theory, especially Mincer's equation for the empirical evidence.

The methodology followed is quantitative, non-experimental, cross-sectional or transversal, specifically correlational or causal, in which the estimation of the ordinary least squares regression was applied in two stages, the first assuming as dependent variable the logarithm of income and as independent variables the educational level achieved, experience and gender, and in a second stage adding to the proposed model the tics variable referring to the use or not of the computer and if this fact impacts on income.

The results obtained have confirmed the hypotheses of the research, the level of education has a positive impact on income by 11% for each additional year of study, the more experience the higher the positive impact on income by 7%, but the fact of being a woman has a differentiated impact, with women earning 39.8% less than men.

In the second stage of the conceptual model, it is concluded that if education increases by one year, the impact on income is an increase of 9%; more experience has an impact on income of 8%; being a woman, income is 41% lower than what men receive; and if they use ICTs, this has a positive impact on their income, 40%, than if they do not use them.

Key words: income, experience, gender, tics, human capital theory.

I. INTRODUCCIÓN

El estudiar en la universidad para la mayoría de los jóvenes en el Perú, significa que a futuro puedan tener acceso a un puesto de trabajo, y a la posibilidad de generarse un autoempleo o un emprendimiento, con el que esperan obtener un sueldo o un ingreso estable que les permita crecer y desarrollar en la vida. Sólo es posible entender esta visión de estos jóvenes, a partir de los fundamentos de la educación, especialmente de la docencia universitaria, investigando la relación que pudiera haber entre el nivel educativo universitario logrado, la experiencia de trabajo, y la situación de género que en una sociedad como la peruana influye, y su relación con los ingresos económicos que se obtenga a futuro.

Una explicación a lo planteado, es pensar que la educación, en términos prácticos, es invertir en uno mismo, esto es generar un mayor valor en nosotros mismos, la idea es entender que somos un capital humano y al invertir en nosotros mismos ese capital se valora más y por ende ese mayor valor en el mercado laboral generará un retorno mayor, quiere decir un mayor ingreso, ese es el planteamiento central de la teoría del capital humano (Becker, Human Capital, A Theoretical and Empirical Analysis, With Special Reference To Education, 1975) con su trabajo pionero en los Estados Unidos y que a partir de él, se han generado, alrededor del mundo, muchas investigaciones explicando la relación directa entre el nivel educativo y el ingreso de las personas tal como reseña (Montenegro & Patrinos, 2014) en el que refiere a estimaciones que se han realizado en 139 países, de las que podemos rescatar de todas ellas la metodología usada, que es el uso de la ecuación de Mincer que evidencia la relación entre el nivel educativo, entre otras variables, con los ingresos.

En el ámbito latinoamericano, también se han realizado este tipo de investigaciones y con esta metodología, tal es el caso del artículo de Galassi y Andrada, presentado en las X Jornadas Argentinas de Estudios de población (Galassi & Andrada, 2009) en el que se explora, con evidencia empírica, la relación entre el salario y la educación en seis zonas geográficas en la Argentina, o en trabajos de investigación con ciertas variantes, al modelo original, como este que explora el efecto de las decisiones laborales conjuntas sobre la estimación de

ecuaciones de Mincer en parejas, esto es, explora la relación de la decisión de trabajar y el salario (Alejo & Funes Leal, 2021). Otra variante de este tipo de trabajos de investigación es la que propone una metodología para descomponer el efecto marginal de la educación sobre la desigualdad salarial, en la relación entre el salario (ingresos) y el nivel educativo (Alejo, 2018).

En el Perú, son varias las investigaciones bajo esta línea de trabajo, y entre las más importantes, está por ejemplo la de Yamada, Lavado y Oviedo, quienes, entre uno de sus principales hallazgos, encuentran que el 62% de las combinaciones universidad-carrera muestran retornos económicos positivos y un 4% de los binomios instituto-carrera muestran retornos positivos (Yamada, Lavado, & Oviedo, La evidencia de rendimientos de la educación superior a partir de "Ponte en Carrera", 2016), a la luz del observatorio "Ponte en Carrera" que es una página web de información pública sobre la educación superior y la demanda laboral, creada en el 2015, a través de una alianza estratégica entre el Ministerio de Trabajo y Promoción del Empleo, el Ministerio de Educación e IPAE Acción Empresarial, para implementar un observatorio de Educación y Empleo (Mtpe, Minedu e Ipaee, 2015).

En enero del 2020, la Superintendencia Nacional de Educación Superior (SUNEDU) emitió el segundo informe sobre la realidad universitaria en el Perú, informe a nivel descriptivo, en base a cinco aspectos, la oferta universitaria, las planas docentes, la investigación, las brechas de género y las condiciones laborales de los egresados universitarios, en este último aspecto, relevante para nuestro propósito, se plantearon interesantes hallazgos que reflejan la situación laboral de los egresados universitarios en términos de sus condiciones de trabajo y sus retornos económicos, por ejemplo el de que los egresados universitarios ganan 73.7% más que sus pares sin educación superior y que el salario de las mujeres egresadas, en promedio, es un 22.1% menor al compararlo con el de los hombres, entre otros, ¿qué metodología se usó para obtener estos hallazgos?, en el anexo metodológico de ese informe está la respuesta, se usaron como fuente de información, entre otras, la Encuesta Nacional de Hogares (ENAHOG), los fundamentos que propuso la ecuación de Mincer y como en todos los artículos e investigaciones revisadas, la metodología de los modelos estadísticos avanzados,

para obtener la estimación, en base a los modelos de regresión lineal, entre el nivel educativo, en este caso de educación superior y los retornos laborales que ello implica, los ingresos (Superintendencia Nacional de Educación Superior Universitaria, 2020). En la actualidad, tras el embate de la pandemia del COVID19, se han evidenciado las nuevas necesidades del mundo de la era del conocimiento, en el que hace rato vivimos, con la urgencia de manejar las tecnologías de la información y comunicación (TICs), en un mundo en el que predomina el uso del internet, del teléfono móvil, de la computadora o laptop, del manejo de las redes sociales y otros, y que tienen un rol central en todas nuestras actividades personales y profesionales, lo que nos obliga a conocerlas y manejarlas obligatoriamente porque si no, estaremos desconectados de la vida económica, social y cultural, tal como informa una organización emblemática dedicada a la Educación, Ciencia y Cultura (Unesco, 2016).

En este sentido, consideramos importante revisar la probable relación entre las TICs y los ingresos, destacando un trabajo de investigación que analiza el uso de estas tecnologías en el trabajo y su incidencia sobre los salarios de manera novedosa para Argentina, en el que incorporan como variable independiente el uso de equipos informáticos, y el uso de software y sus impactos en los salarios, encontrando que los trabajadores con uso de PC presentan un diferencial que va entre el 14% y 19% de sus salarios con respecto a los que no utilizan, y en mayor medida es el diferencial para el caso de los desarrolladores de softwares (Brassiolo, Nahiriñak, & Ruffo, 2006). Otro caso, en esta misma línea, es el trabajo de investigación que prueba el uso de las tecnologías de la información en el trabajo y sus implicancias en las diferencias salariales en México, que muestra hallazgos relacionados a la existencia de disparidad salarial a favor de los trabajadores con habilidades informáticas, así como en la experiencia en el uso de las computadoras y su implicancia en los salarios (Torres García & Ochoa Adame, 2018). Para una realidad como la peruana, lo primero que se ha considerado, es revisar algunos datos clave como, por ejemplo, los datos relacionados a la tenencia de la infraestructura básica de las TICs por parte de los egresados universitarios (Instituto Nacional de Estadística e Informática, 2015), y esto es lo que nos refiere la última encuesta realizada a estos egresados:

Tabla 1

Tenencia de infraestructura básica de las TICs de egresados universitarios

Comunicación e informática	Abs	%
Telefonía móvil	176,803	90.1
Telefonía fija	119,993	61.1
Computadora (PC)	151,625	77.3
Laptop	146,437	74.6
Tablet	73,517	37.5
Total	196,240	

Nota: INEI – Encuesta N. de Egresados y Univ. 2014

Observamos que mayoritariamente contaban con teléfono móvil, computadora y laptop al momento en el que se les realizó la encuesta, ahora el tema es el uso que le daban a estos equipos básicos y su uso en el internet, la encuesta nos señala que el 98.5% utilizaban internet y sólo el 1.5% no lo hacía, pero la clave es saber la siguiente información:

Tabla 2

Uso de los equipos y el internet de los Egresados universitarios

Uso del internet	Abs	%
Comunicación (email, chat)	178,880	92.5
Comercio electrónico	77,215	39.9
Operaciones banca	79,935	41.3
Curso/educación	121,035	62.6
Búsqueda de trabajo	101,585	52.5
Actividades de investigación y desarrollo	157,602	81.5
Revisión de noticias	165,410	85.5
Entretenimiento	161,722	83.6
Otro	2,555	1.3

Nota: INEI – Encuesta N. de Egresados y Univ. 2014

Y en esta información encontramos que le dan un uso muy limitado, principalmente sólo para comunicarse, para estar al día en las noticias y para el entretenimiento, y muy limitado a lo que es el uso en una práctica laboral cotidiana en la actualidad, las operaciones de comercio electrónico y las operaciones relacionadas a las transacciones bancarias y comerciales.

Este hecho ameritó se indague, sobre si los egresados se han formado con las innovaciones pedagógicas aportadas con el uso de las TICs y de acuerdo con los requerimientos que la demanda laboral exige en la actualidad, resulta que al revisar la percepción que los egresados tenían respecto al desarrollo de sus habilidades y competencias, respecto al desempeño de sus docentes, y respecto a sus universidades, en sólo lo pertinente a los alcances de la presente investigación, se encontró que con respecto a lo primero, dan un mayor valor al trabajo en equipo, al tener liderazgo y al trabajo bajo presión, que respecto a utilizar herramientas informáticas básicas y al uso de software en sus carreras. Con respecto a sus docentes, se encontró que son más exigentes en sus apreciaciones respecto a la calidad de los profesores y al grado de actualización de estos, pero su percepción es crítica, especialmente con respecto a los docentes de las universidades públicas versus los docentes de las universidades privadas, específicamente sobre el grado de actualización de los conocimientos de sus docentes. Y finalmente, la percepción de los egresados con respecto a sus universidades, en el entendido de que el uso de las TICs ha influenciado los cambios en los procesos educativos, ellos manifiestan que el 71.3% de las universidades, a la fecha de la encuesta, contaban con aulas virtuales, el 80.5% en las privadas y sólo el 52.5% en las públicas y todos, manifiestan, que estaban orientadas especialmente a las matrículas, a los registros académicos y al trámite documentario (Instituto Nacional de Estadística e Informática, 2015). Esta problemática y la revisión de los diversos trabajos de investigación que se han realizado a nivel mundial, regional y local, significan no solo un reto académico para evidenciar empíricamente la relación práctica entre las variables independientes, logro del nivel educativo superior, la experiencia, y la variable género y su correlación causal y explicativa en la variable ingreso, sino un aporte en materia de innovación pedagógica por la comprobación empírica del rol de las TICs en el ingreso, y más aún si esto último, no se ha investigado, bajo esta perspectiva, en nuestro país a la fecha.

Esta problemática y su perspectiva de análisis, planteada en el párrafo precedente, nos permite precisar la pregunta de investigación en dos etapas, en una primera etapa, ¿cuál es el impacto, del nivel educativo logrado, la experiencia que tienen, y el género, de los egresados universitarios, en los ingresos que

perciben en su actividad laboral?, y en una segunda etapa, ¿cuál es el impacto, del nivel educativo logrado, el manejo de las TICs, la experiencia que tienen, y el género, de los egresados universitarios, en los ingresos que perciben en su actividad laboral?.

No está demás manifestar, que los hallazgos de esta investigación contribuirán con el enriquecimiento del conocimiento del aporte de las TICs en los procesos de innovación pedagógica que se manifiestan en términos absolutamente prácticos al generar ingresos diferenciados como consecuencia de haber logrado un nivel educativo universitario y al estar al día en el manejo de las herramientas que las TICs implican en el trabajo laboral, sino que este hecho se convierte en un mensaje claro y directo a toda la juventud que va a entrar próximamente al mercado laboral, quieres mejores ingresos, estudia y capacítate permanentemente. Y también, con la presente investigación, se contribuirá a resolver un problema surgido como consecuencia de la Covid 19, el tener que estar listos para cualquier eventualidad respecto a la probabilidad de tener que usar el teletrabajo en perspectiva, por cualquier circunstancia que nos obligue al necesario distanciamiento social.

Por todo ello en esta investigación el objetivo general, en una primera etapa, es determinar el impacto, del nivel educativo logrado, la experiencia que tienen, y el género, de los egresados universitarios, en el nivel de ingresos que perciben en su actividad laboral; y en una segunda etapa, determinar el impacto, del nivel educativo logrado, el manejo de las TICs, la experiencia que tienen, y el género, de los egresados universitarios, en el nivel de ingresos que perciben en su actividad laboral. Por tanto, la hipótesis general del presente trabajo de investigación que postulamos es, que en una primera etapa, existe un impacto positivo, del nivel educativo logrado y la experiencia de los egresados universitarios, en el nivel de ingresos de estos en su actividad laboral, y un impacto diferenciado entre el género de los egresados universitarios con el nivel de ingresos de estos en su actividad laboral, y en una segunda etapa existe un impacto positivo, del nivel educativo logrado, el manejo de la TICs y la experiencia, de los egresados universitarios, en el nivel de ingresos de estos en su actividad laboral, y un impacto diferenciado entre

el género de los egresados universitarios con el nivel de ingresos de estos en su actividad laboral.

II. MARCO TEÓRICO

En este capítulo, presentamos una síntesis de los antecedentes internacionales y nacionales investigados, las teorías y los enfoques conceptuales, y la epistemología de la problemática investigada.

Con respecto a los antecedentes internacionales nos referiremos a (Montenegro & Patrinos, 2014) quienes, en su investigación relacionada a las estimaciones comparadas del rendimiento de los niveles de educación en todo el mundo, presentaron estimaciones comparables con datos de 139 países en el mundo, empleando 819 encuestas de hogar, debidamente armonizados. La idea en esta investigación era demostrar en términos empíricos la relación entre los retornos de la educación y la experiencia lograda, con los mayores salarios, esto es, la búsqueda de mejores niveles educativos para obtener mejores retornos en la concepción de que los rendimientos de los niveles educativos y de la experiencia son indicadores de la productividad de las personas que conllevan a obtener mejores salarios. Usaron la ecuación de Mincer, en la que se relaciona al logaritmo natural del salario, con el número de años de educación logrado, con la experiencia laboral en número de años al cuadrado y con un término de perturbación. Utilizando bases de datos generados en el Banco Mundial, y con los métodos de mínimos cuadrados se lograron obtener resultados interesantes como lo establece que es en la estimación el modelo de Mincer, que hay una asociación positiva entre nivel educativo y salarios, que hay una relación positiva entre el salario y la experiencia, pero que en el largo plazo hay unos rendimientos decrecientes entre la experiencia y los salarios.

Otro artículo relevante a considerar es el de (Alshyab, Sandri, & Abu-Lila, 2018) quienes investigaron las diferencias que se presentan en el rendimiento de la educación, en especial a partir del análisis de género y por sectores, público y privado, investigación efectuada para Jordania, país ubicado en la región de oriente en Asia. Esta investigación analiza la relación entre educación e ingresos en Jordania en relación con el género y el sector de empleo, sector público y sector privado. Los resultados que muestran están en que los retornos de la educación para los hombres empleados son mayores que los retornos para las mujeres

empleadas, esto es que la prima salarial de un año adicional de escolaridad (nivel educativo) es de 6.8% para hombres en tanto que para mujeres es 5.5%, otro resultado interesante es que las mujeres obtienen, en promedio, 74% de lo que ganan los hombres. Para el logro de estos resultados utilizan el modelo de la ecuación de Mincer, pero con alguna modificación a su versión original, y con la utilización como método, las técnicas estadísticas de correlación avanzada llamada modelos de efectos fijos sobre un conjunto de datos próximo a los llamados data panel que cubren el periodo entre el 2000 y el 2015.

(Ons Cappa, Sánchez Ollero, & García Pozo, 2020) en una investigación sobre las diferencias que se producen en los rendimientos del capital humano cuando se considera el género en el sector de turismo en España, emplearon una versión extendida de la función salarial de Mincer con datos referidos al año 2014 procedentes de la Encuesta de Estructura Salarial (EES) que periódicamente se realiza en ese país. Como variable dependiente consideraron al salario bruto por hora expresado en logaritmos, en tanto que como variables independientes utilizaron al conjunto de variables definidos como capital humano y como aporte de esta investigación consideraron otras variables referidas a determinadas características personales y laborales del trabajador/a cómo, los años promedio de estudios y la experiencia previa. En el caso de los años promedio de estudios los clasificaron en analfabetos o sin estudios, con estudios primarios, los que completaron el primer ciclo de educación secundaria, los que tienen formación profesional de primer grado, o de segundo grado, los que tienen diplomado universitario, los que tienen licenciatura y los que tiene post grado, en tanto que para la variable experiencia la definieron como la diferencia entre la edad del individuo observado y la edad a la que declara haber comenzado a trabajar, limitada o restringida por la edad de jubilación y los años de estudio. Sus principales resultados muestran que los rendimientos de la educación son estadísticamente significativos y positivos para las mujeres y para los hombres. Los resultados respecto a la experiencia señalan que para el caso de los hombres no es relevante el efecto en los salarios, pero si para el caso de las mujeres, aunque con rendimientos negativos.

Esta otra investigación de (Villarreal Peralta, 2018) estudia la relación de los rendimientos del nivel educativo logrado con los salarios que obtienen en su actividad laboral los trabajadores en México, en dicha investigación se ha obtenido que en promedio un año de escolaridad (nivel educativo) hace crecer en 9% el salario de los trabajadores para el periodo de estudio, 1995 – 2012, y ¿cómo se obtuvo este resultado?, investigaron la evolución histórica de los rendimientos educativos para México, evaluando el impacto del nivel educativo sobre los ingresos de los trabajadores mexicanos en dicho periodo, y lo hicieron en base a tres estimaciones, la primera a través de una ecuación de Mincer tradicional que medía el incremento salarial del trabajador por un aumento de un año de escolaridad, la segunda con la metodología de variables instrumentales con las que controlaban la endogeneidad de la variable escolaridad a través del cambio legislativo sobre los años de escolaridad obligatoria y la tercera por niveles educativos con los que han medido la prima salarial de acuerdo a su avance. Ahora bien para estudiar la evolución histórica de los rendimientos educativos han utilizado el modelo de Mincer, en el que la variable dependiente es el logaritmo natural de los salarios individuales por hora trabajada, medida de la productividad laboral, y como variables independientes están una variable que es lo aproximado a la educación formal expresada mediante el número de años de estudio finalizados por la población, otra la experiencia laboral y la experiencia laboral al cuadrado, otra variable correspondiente al sexo del individuo y la perturbación aleatoria. Utilizaron como método de estimación a los mínimos cuadrados ordinarios, pero corregido por el uso de variables instrumentales, y esa es la particularidad de este modelo, considerando las leyes de escolaridad obligatoria como un determinante exógeno del nivel educativo alcanzado.

Con este otro artículo avanzamos mucho más, particularmente en la consolidación de nuestro marco conceptual y los fundamentos que requerimos para el desarrollo de nuestra investigación, es el caso del estudio realizado en los jefes de hogar en las familias de Colombia, respecto a los retornos de la educación y su efecto en los ingresos de estas personas, escrito por (Suárez Díaz & Fuentes López, 2020) quienes presentan en esta investigación una estructura conceptual entorno al capital humano y la ecuación de los salarios, con su respectiva evidencia

empírica a través de un análisis de estadística avanzada de la dinámica de salarios de los jefes de hogar en Colombia para los años 2013, 2014 y 2015. Con respecto a la estructura conceptual, presentan el estado del arte sobre la rentabilidad de los salarios y las metodologías utilizadas para su análisis, también desarrollan sistemáticamente el marco teórico de la educación y las tasas de retorno de este tipo de inversiones, y proponen que cuando un individuo se traza metas respecto a aumentar sus ingresos salariales, esto puede lograrlos a través de una inversión en capital humano, en donde la educación, la especialización en su profesión o labor, la salud o la seguridad, entre otros aspectos, constituyen un mayor capital humano y una inversión de la persona sobre si misma que generará retornos, recuperando lo invertido más una ganancia adicional. Para los análisis respecto a la evidencia empírica de los retornos de la educación en el periodo de referencia, utilizaron los microdatos de las Encuestas Nacionales de Calidad de Vida (ENCV) de estos periodos en las siguientes variables, sexo, años cumplidos, parentesco con el jefe de hogar, estado civil, nivel educativo más alto alcanzado y último grado, años de estudios superiores realizados y aprobados, salario mensual y labor que desempeña. Los resultados obtenidos concluyen que se cumple con la teoría del capital humano, toda vez que la probabilidad de tener un mejor retorno de la educación depende tanto de los años de educación y formación académica, como de la experiencia, un año más de educación tiene como efecto la probabilidad de que un jefe de hogar gane mensualmente un mejor salario en Colombia.

En esta otra investigación que estudia en el mercado laboral colombiano, las discriminaciones que se producen como consecuencia o de ser hombre o de ser mujer en las remuneraciones, se hace un análisis de los elementos diferenciadores del salario en Colombia que dependen de las diferencias en la educación, la experiencia, la tenencia y el estado civil, pero de manera especial del género, para el 2018. La metodología usada es de índole histórico descriptivo cuantitativo que en base al modelo de Mincer plantean que las diferencias salariales están dadas por factores tales como la edad y la experiencia laboral, esto señala que tendrá mayor rentabilidad salarial quien más nivel educativo tenga, independientemente de la edad, pero complementariamente a estas diferencias, toman en cuenta la teoría de la discriminación propuesta por Gary Becker, Keneth Arrow, Edmund

Phelps, entre otros. Entendiéndose como discriminación, al rechazo a contratar a una persona con un valor de producto marginal más alto que su costo marginal, esto es la discriminación que está ligada a la productividad, y estadísticamente entendiéndose que la diferencia entre los salarios de las personas se debe a que dentro del mercado existen grupos que tienen características propias y que el mercado las valora más, ósea el mercado no sólo valora a la productividad sino también a las características personales. Producto de la estimación del modelo utilizado en esta investigación, se encontró que los coeficientes obtenidos tienen los signos esperados y son estadísticamente significativos, encontrándose que existe una brecha salarial entre hombres y mujeres, no sólo explicada por las diferencias en productividad sino en otros factores. En concreto, se encontró que en promedio el género femenino, cabeza de hogar en Colombia, devenga salarios menores que sus pares masculinos, no sólo por efecto de las diferencias en los niveles de educación, sino por otros factores vinculados a la propia naturaleza de ser mujeres en esta realidad, vinculadas a diversas obligaciones familiares, a las que enfrentan tales como atender el cuidado de los niños, o el cuidado de los adultos mayores, y con el consiguiente impacto en los hogares completos, menores salarios mayor pobreza (Pérez, Fuentes, & Contreras, 2020).

Una investigación realizada para Colombia, los retornos de la educación universitaria de la población ocupada en dicho país, son evidenciados con la existencia de un bajo porcentaje de ocupados con título universitario a pesar de que, si es rentable realizar una carrera universitaria ya que, al ingresar al mercado laboral, contando con un título de educación superior, se tiene una rentabilidad promedio de 11.2% para hombres y 9.1% para mujeres por sobre quien no tiene. Además, señalan que por cada año adicional de escolaridad (nivel educativo), el nivel de ingresos se eleva en 2.45% en general, pero el tener un título universitario le genera un incremento salarial promedio de 48.7% respecto al que tiene sólo el bachiller y, si se obtiene un título de post grado, el aumento salarial sería de 75.6% respecto del que sólo logro los estudios universitarios. Estos resultados se obtuvieron por que la investigación consideró la ecuación de Mincer como punto de partida, complementada con variables de control que captura los aspectos diferenciadores de los individuos, con lo que la ecuación completa comprendía al

logaritmo natural del ingreso mensual de los individuos, que está en función del número de los años de escolaridad de estas personas, el nivel educativo, la experiencia del individuo, la experiencia del individuo al cuadrado, y además incluyeron algunas variables de control dicotómicas asociadas a los departamentos, asociadas al género, tipos de empleo del individuo, el ser empleado del sector público o del sector privado o a cuenta propia, variables de dominio (urbano, rural) y en relación al tipo de contrato (Gil León, Casas Herrera, & Lemus Vergara, 2020).

Castillo, Da Silva y Pérez (2017) en una investigación realizada para Colombia, estudiaron los retornos salariales, pero incorporando un nuevo método de estimación en su análisis, el método cuantílico, en él estudiaron los retornos salariales de los jóvenes y adultos, comparando la influencia de la educación y la experiencia. Para ello, analizaron algunos postulados de la relación entre educación, experiencia y mercado laboral, parten de la teoría del capital humano como eje central de la relación educación-productividad-salarios, pero incluyen las críticas sostenidas a estos postulados e incluyen en sus análisis a la experiencia como fundamento en la capacitación y determinación salarial de las personas, en base a la teoría del credencialismo, que plantea que la productividad viene dada también por las credenciales que le permiten competir a las personas por un puesto de trabajo. También plantean, en base a la teoría institucionalista que postula que lo más importante es la formación adquirida en el puesto de trabajo, a la experiencia. También consideran, en base a la teoría del mercado dual del trabajo que señala que existen dos tramos en el mercado laboral, uno el primario, que tiene altas remuneraciones salariales con posibilidades de ascensos y estabilidad, y otro el secundario con menores remuneraciones salariales, menores posibilidades de ascenso y más inestable. Consideran también las nuevas teorías del mercado laboral como el de los mercados transicionales en los que se introduce el concepto del ciclo de vida de las personas y los problemas de inserción laboral. Para su análisis empírico usan el modelo de Mincer, inicialmente en base a los mínimos cuadrados ordinarios y luego en base a la regresión cuantílica como desarrolla Vicéns y Sánchez (2012), quienes señalan que este método es un tipo de análisis de regresión que a diferencia del método de mínimos cuadrados ordinarios, arroja

estimaciones que se aproximan a la esperanza condicional de la variable respuesta, dados ciertos valores de predicción, y que tiene por objeto estimar la mediana condicional u otros cuantiles de la variable respuesta. Con este método y utilizando las variables de la Gran Encuesta Integrada de Hogares para el año 2015, realizan estimaciones empíricas primero con los mínimos cuadrados ordinarios y luego con la regresión cuantílica considerando como variable dependiente al ingreso laboral, y como variables independientes a la educación, medida en años de escolaridad, a la experiencia, medida como la suma de la duración en años del empleo anterior y actual, a la variable género, considerada como una variable dummy (mujer u hombre), la variable contrato (tener contrato o no tener contrato), asimismo otras variables como departamento donde trabaja o zona geográfica que refiere, y por rama de actividad económica.

En esta otra investigación, realizada para Ecuador por Lucero (2019) se investiga la relación entre educación e ingresos, establecida por la teoría del capital humano, en la que propone demostrar que, tener más años de educación formal, contar con un perfil de mano de obra que se está formando, tener calidad en la educación recibida y fomentar la universalización de la educación fortaleciendo sus dinámicas laborales, impedirían que los rendimientos educativos sean negativos. Utilizó modelos de corte transversal para demostrar su hipótesis de trabajo para cada año en el periodo 2007 – 2017, con datos de la Encuesta Nacional de Empleo, Subempleo y Desempleo (ENEMDU) que realiza el Instituto Nacional de Estadística y Censos (INEC) del Ecuador y el modelo Mincer. Los resultados confirmaron que mayores niveles de educación generan ventajas en la obtención de mejores ingresos, pero esta ventaja en el tiempo presenta una tendencia decreciente, lo que a partir de ello resultan varios desafíos que plantear al diseño de la política educativa que tiendan a mejorar a la calidad y a su pertinencia para el mercado laboral en el Ecuador.

También referimos a esta otra investigación realizada por Gálvez y Valdés (2019) quienes estudian el retorno de la educación superior en Chile, en una muestra de trabajadores asalariados obtenida de la Encuesta de Caracterización Socioeconómica (CASEN) del 2015 realizada por el Ministerio de Desarrollo Social de Chile, estimando el retorno de la educación en aquellos trabajadores formados

en un centro de formación técnica (CFT), en institutos profesionales (IP), o en una universidad estatal o en aquellas universidades históricas que pertenecen al Consejo de Rectores de Chile (CRUCH), también analizaron los retornos diferenciales por género de aquellas personas con educación superior. El modelo que utilizaron inicialmente, fue el de Mincer, en el que la variable dependiente fue el logaritmo del salario por hora, y como variables independientes tuvieron a los años de estudio del individuo, la variable que indica la experiencia laboral medida en años del trabajador como variable proxy, y la variable que representa al error aleatorio. Sin embargo, posteriormente modificaron este modelo incluyendo variables de control dicotómicas especificando si el individuo materia de estudio, estudió en un CFT, o si estudio en IP, si estudió en una universidad no tradicional ó en una universidad tradicional, en este modelo también incorporan variables exógenas relacionadas a la nacionalidad, género, origen étnico, zona y región de residencia del individuo. Con estos fundamentos y estos modelos, obtuvieron resultados que señalan que la educación superior en Chile tiene un retorno mayor que en otros niveles de educación más bajos, que entre las instituciones formativas, las universidades tradicionales (CRUCH) tienen el retorno de la educación más elevado, que a su vez las universidades no tradicionales tienen los retornos mayores que los IP y los CFT, y que las mujeres y las personas que pertenecen a alguna etnia tienen menores retornos, y finalmente que las mujeres tienen menores retornos comparado a el de los hombres.

Gómez (2018) en su trabajo de investigación analiza empíricamente los retornos de la educación y los premios por calificación en asalariados ocupados en Argentina, entre 2003 y 2014, en base a la utilización de los microdatos de la Encuesta Permanente de Hogares (EPHC) con una periodicidad trimestral, encuesta que se hace sobre las condiciones de vida de la población, centrándose básicamente en el mercado laboral. Utilizó tres especificaciones alternas de la ecuación de Mincer, con las variables educación y experiencia, pero adicionándole el tipo de formación, el género, la región geográfica, la formalidad del empleo, la rama de actividad, y la calificación del trabajo. La primera alternativa fue estimar en base a los mínimos cuadrados ordinarios, la segunda, en base a regresiones cuantílicas incondicionales y la tercera alternativa es utilizando regresiones

cuantílicas condicionales. Con esta propuesta la autora evaluó, qué estimación alterna resulta empíricamente la más apropiada para el caso argentino, los resultados muestran que el modelo de Poisson de máxima verosimilitud (las regresiones cuantílicas, segunda y tercera alternativa) aplicada al enfoque de la ecuación de Mincer son estadísticamente más consistentes y robustas en los retornos a los atributos de los trabajadores.

Para incorporar la variable TICs en esta investigación, en párrafos del capítulo anterior se ha citado a la investigación de Brassiolo, Nahirñak y Ruffo (2006), en el que se incorpora la variable tecnología informática, medida en dos dimensiones, una que señala si usa en el puesto de trabajo una computadora y otra si en el puesto de trabajo usan software, adicional a las otras variables referidas a la educación, antigüedad en el cargo, experiencia laboral, género, categoría ocupacional, calificación del puesto, entre otras más. Los resultados en esta investigación hecha para la Argentina, señalan que este hecho produce un diferencial salarial que va entre el 14% y el 19% en aquellos trabajadores que en su labor usan la computadora, y el diferencial es mucho mayor si usan, en su labor software. También se citó la investigación de Torres y Ochoa (2018), quienes estudiaron las implicancias en el diferencial salarial en México como efecto del uso de las TICs, para ello desarrollaron una regresión por el método de mínimos cuadrados ordinarios considerando como variable dependiente al logaritmo natural del salario por hora y como variables independientes a la escolaridad (nivel educativo), sexo, estado civil, experiencia, horas trabajadas, ciudad, rama, tamaño, ocupación, y la variable usa-TIC considerada como una variable dicotómica, esto es, toma valores de cero y uno, será uno si usan estas tecnologías en el lugar de trabajo y cero si no usa, también consideraron como variable si se usa computadora de forma diaria y si se tiene experiencia en el uso de estas herramientas. Los resultados obtenidos en esta investigación señalan que los que utilizan las TIC tienen un salario mayor en 2.8%, si este conocimiento es usado en el lugar de trabajo tienen una prima salarial de alrededor del 13%, si usan a diario la computadora tiene un diferencial salarial de 8% y si tienen experiencia en el uso de las computadoras obtienen un premio salarial del 11.6%.

Adicional a estas referencias se considera también la investigación de Félix y Torres (2018), quienes investigan el efecto del uso de la computadora en el trabajo en los salarios en México, para ello utilizan dos métodos, el primero es el uso de las técnicas estadísticas avanzadas de correlación de variables y el otro método es el uso de las técnicas de emparejamiento para identificar las diferencias salariales entre los usuarios y no usuarios de la computadora en su trabajo. Para el primer método consideran como variable dependiente al logaritmo natural del salario por hora de los trabajadores, y como variables independientes a los años de escolaridad (nivel educativo), a la variable binaria género, estado civil, experiencia, a la variable binaria con valor igual a 1 si se utiliza computadora en el trabajo y 0 si no se utiliza, y a la perturbación aleatoria. Consideran algunas variaciones a este modelo, incorporando gradualmente otras variables como una referida al uso diario de la computadora en el trabajo y otra que señala el uso semanal de esta herramienta. Para el uso del segundo método utilizan dos técnicas con las que cuantifican las posibles diferencias en base a un método de emparejamiento, un individuo y su par más cercano, y el otro utilizando un sistema de puntuación por propensión, ambas técnicas miden una puntuación si se utiliza o no computadora en el trabajo y su efecto en los ingresos. Los resultados con ambos métodos señalan que se producen diferencias salariales por la utilización de las computadoras en el trabajo, en el rango entre 17.2% a 25% para el año 2006, en tanto que en el rango entre 12.7% a 20.2% para el año 2014, resultados con los que les permite concluir que en México existe una diferencia salarial por el uso de la computadora en el trabajo.

Adicional a las investigaciones hechas y referidas para el Perú en el capítulo anterior, se cita el trabajo de investigación de Ventura (2012) que trata sobre los retornos de la educación en el Perú y su probable heterogeneidad a nivel regional. En esta investigación se midió los retornos de la educación a partir de una muestra de personas asalariadas en base a un modelo en el que la variable dependiente es el logaritmo natural del salario por hora del individuo, en una determinada región y en un determinado tiempo, y como variables independientes se tiene a los años de educación del individuo, en la región tal y en un determinado tiempo, otra variable independiente está referida a las características personales del individuo como

edad, sexo, tipo de centro de estudios y experiencia y la otra variable es el gasto público agregado y desagregado por funciones. Sus principales resultados muestran que en todas sus estimaciones, en base a la información de la Encuesta Nacional de Hogares del 2004 al 2010, se obtiene retornos a la educación básica de 8.1%, retornos para los que culminaron la educación secundaria de 7.6%, retornos para los que culminaron la educación superior de 10.5%, los retornos por género son de 30.9%, por experiencia 1.9% y por edad 1.1%, todos estos resultados son estadísticamente significativos al 5%. En cuanto a la variable gasto público, se ha obtenido, en primer lugar, que las correlaciones muestran una relación directa entre el nivel de gasto público regional y los salarios, en segundo lugar, se ha obtenido como resultado que 1% más de gasto público está asociado a un aumento en los salarios de 15%, y en tercer lugar se ha obtenido como resultado que el gasto público regional afecta indirectamente el nivel de los salarios. Con estos resultados esta tesis concluye que ha demostrado la existencia de heterogeneidad en los retornos a la educación a nivel regional y cómo estas son influenciadas por el gasto público ejecutado en ellas.

En relación a las bases teóricas y enfoques conceptuales en esta investigación, el fundamento central es el enfoque económico de la educación que esencialmente contribuye en fortalecer, la relación existente entre la educación y el mundo económico tal como señalan Del Campo y Salcines (2008), y en ella la teoría del capital humano que está basada en los aportes pioneros de Schultz, Becker y Mincer, el primero quien sostiene que el adquirir el capital humano, es un proceso conducente a acceder conocimientos y habilidades, que incrementan la productividad y las ganancias, y para ello para acceder a esos conocimientos y habilidades se tiene que invertir para lograr este propósito, este hecho va a generar que ese esfuerzo proporcione una tasa positiva de rendimiento, lo que explicaría que se producirán diferencias en la cantidad de inversión humana que a su vez explican las diferencias en los ingresos (Schultz, 1999). En tanto Becker plantea, en su libro “El capital humano ...”, un análisis de los efectos que se tienen sobre las retribuciones y tasas de rendimiento de la inversión en educación que realizan las personas, entendiendo al capital humano como los conocimientos, las habilidades, la buena salud y sus valores (puntualidad y honestidad) que tienen las personas

como consecuencia de la inversión que realizan a lo largo de su vida, justamente en educación, en salud, en una permanente capacitación, y en todo aquello que lo revaloriza como capital. Las personas, incide, invierten a lo largo de su vida en ellas mismas, ya sea dedicándole tiempo o dinero, para educarse, formarse y también adquirir experiencia, esto les permite, afirma, contar con conocimientos, habilidades, nuevas capacidades con las que podrán influir en sus vidas económicas y por ende en sus ingresos futuros, a partir de la mejora en sus productividades que afectan positivamente en el crecimiento económico y por ende en el incremento de los ingresos. Las personas cuando invierten lo hacen de manera directa al asumir los costos de las matrículas, pasajes, equipos, libros, y otros más, y de manera indirecta al no obtener ingresos que pudieran haber generado y que no lo han hecho por justamente haber dedicado su tiempo en adquirir los conocimientos y las capacidades para revalorarse como capital. Pero el autor incide en señalar también, que como en toda inversión, sólo la inversión en capital humano sería rentable, si los beneficios que esto generaría son mayores a los costos en los que se incurrirá y eso es lo que determina la tasa de retorno, dato clave para la teoría del capital humano y para la relación esencial entre el sistema educativo y el mercado de trabajo (Becker, 1984). Otro aporte fundamental a la teoría del capital humano es la de Mincer, más conocida como la “ecuación de ingresos de Mincer”, quien plantea la relación existente entre los salarios o ingresos con los años de escolaridad (lo planteamos como el nivel educativo logrado) y la experiencia adquirida. La formulación de esta ecuación tal como está planteada en la investigación de Yamada y Castro (2010):

$$\text{Ln } Y(s, x) = \beta_0 + \beta_1 s + \beta_2 x + \beta_3 x^2 + \varepsilon \quad (1)$$

Donde:

Lg Y(s, x) = logaritmo de los ingresos laborales de una persona con

s = años de educación

x = experiencia laboral potencial

ε = error no sistemático

- β_0 = ingreso laboral de una persona sin estudios ni experiencia
- β_1 = retorno porcentual que brinda un año de educación adicional
- β_2 = mide el efecto de la experiencia sobre los ingresos
- β_3 = capta la concavidad de los perfiles de ingreso con respecto a la experiencia

Este es el modelo teórico básico que fundamenta esta investigación, pero con dos variantes, una agregándole la variable género y otra agregando la variable TICs, cuyo desarrollo lo planteamos en el capítulo metodológico.

III. METODOLOGÍA

En este capítulo se presenta la metodología que se utilizó en este trabajo de investigación y que comprende los siguientes aspectos, el tipo y diseño de investigación, las variables y su operacionalización, la definición de la población y muestra, las técnicas e instrumentos de recolección de datos, los procedimientos que se utilizaron, el método de análisis de los datos y los aspectos éticos.

3.1 Tipo y diseño de investigación

La ruta de la presente investigación es cuantitativa y el tipo de diseño es no experimental, de tipo transeccional o transversal, específicamente de tipo correlacional o causal (Hernández-Sampieri & Mendoza Torres, 2018), con datos del 2019.

Se consideró esta metodología porque en esta investigación no se manipuló a ninguna variable, esto es, no se varió de forma intencional a las variables independientes para observar sus efectos sobre otras variables, lo presentamos tal cual, lo que se hizo es observar y analizar el impacto de las variables independientes sobre la variable dependiente en un momento del tiempo, esto es para el año 2019 porque es un año, cuya data es no atípica.

Se planteó en la presente investigación para el trabajo empírico, el desarrollo del modelo teórico en dos etapas utilizando la metodología de los mínimos cuadrados ordinarios en modelos con análisis de regresión múltiple (Wooldridge, 2010), esto significa realizar regresiones entre una variable dependiente y varias variables independientes lo que nos permitirá, utilizando la ecuación de Mincer, para medir y comprobar los impactos del nivel de educación, la experiencia, la variable género y la variables TICs en los ingresos de las personas, específicamente de los egresados universitarios, y para esto, inicialmente se plantea el siguiente modelo econométrico:

$$\log \text{Ing} = \beta_0 + \beta_1 \text{Educ} + \beta_2 \text{Edad} + \beta_3 \text{Edad}^2 + \beta_4 \text{Sexo} + \varepsilon \quad (2)$$

En el que $\log \text{Ing}$ es el logaritmo de los ingresos, Educ es la variable que mide el nivel educativo alcanzado, Edad es la variable proxy que representa la experiencia, Sexo es la variable género y el término de perturbación.

Luego en una segunda etapa al modelo econométrico planteado lo ampliamos agregándole la variable TICs tal como se define a continuación:

$$\log \text{Ing} = \beta_0 + \beta_1 \text{Educ} + \beta_2 \text{Edad} + \beta_3 \text{Edad}^2 + \beta_4 \text{Sexo} + \beta_5 \text{TICs} + \varepsilon \quad (3)$$

Con estos dos modelos econométricos (2) y (3) se realizó la estimación de regresión con los datos que nos proporciona la Encuesta Nacional de Hogares (ENAH) para el año 2019 proporcionada por el Instituto de Estadística e Informática (INEI, 2022) con los que se analizó los resultados y la evidencia empírica de las hipótesis de trabajo de esta investigación.

3.2 Variables y operacionalización

La variable dependiente es el ingreso, que conceptualmente se entiende como la “remuneración monetaria o en especie que recibe un trabajador por sus prestaciones en la producción de bienes y servicios dentro de una empresa” (BCRP, 2022). Se debe recordar que dicha remuneración se pacta en forma bilateral, esto es entre la empresa y el trabajador o de forma colectiva, quiere decir a partir de un convenio colectivo que usualmente se suscribe entre el sindicato y la empresa.

En términos operativos, la variable ingreso la expresaremos en términos de variaciones porcentuales o logaritmos, bajo el enfoque de la teoría del capital humano que entiende como los “ingresos que denota a los salarios, tomados como ingresos mensuales” tal como plantea Suárez y Fuentes (2020) y expresados como los ingresos totales percibidos en un determinado periodo, más los pagos percibidos en especie, más los ingresos generados si tuviese labor independiente y más el valor de los productos que utiliza para su consumo.

La única dimensión de esta variable es el logaritmo de los ingresos, y su único indicador es el $\text{Log}(\text{Ing})$, variable cuya escala de medición es ser una variable cuantitativa y expresada en términos de una escala de razón.

Tabla 3

Variable dependiente Log (Ing), dimensión e indicador

Variable	Dimensión	Indicador
Ingreso, es la remuneración monetaria o en especie que recibe un trabajador por sus prestaciones en la producción de bienes y servicios dentro de una empresa.	Logaritmo de los ingresos	Log (Ing)

Entre las variables independientes se tiene al nivel educativo, cuya definición se basa en el concepto del “proceso de formación del individuo que genera un aumento en la capacidad productiva” concepto que refiere (Modrego Rico, 1989) y que cita (Del Campo Villares & Salcines Cristal, 2008) en su trabajo de investigación.

La definición operacional de esta variable es el número de años invertidos en el proceso de su formación educativa que comprende la formación básica regular y la formación superior.

La dimensión es los años de educación logrados, cuyo indicador es el número de años de educación Educ y está expresada en términos cuantitativos y de una escala de razón.

Tabla 4

Variable independiente Educ, dimensión e indicador

Variable	Dimensión	Indicador
Nivel educativo cuyo concepto es el proceso de formación del individuo que genera un aumento en la capacidad productiva.	Años de educación	Número de años de educación

La otra variable independiente es experiencia, variable que se enfoca bajo el concepto de que es la medida que aumenta el nivel de entrenamiento y especialización que hay en una empresa que aumentará también las diferencias salariales entre el grueso de la población (Mincer, 1981).

Como definición operacional se tiene que para esta variable, que indica la experiencia laboral del individuo, se consideró a los años del trabajador como una variable proxy, por la falta de información en las estadísticas que se ha utilizado,

este procedimiento de utilizar variables proxy es bastante regular en aquellas investigaciones que carecen de información sobre un detalle particular en su investigación por lo que, al revisar los trabajos que sirvieron de antecedentes a esta investigación, se encontraron, entre otros, a la investigación de Gálvez, Castex y Valdés (2019) para un caso en Chile.

La única dimensión considerada en esta variable es los años cumplidos, cabe resaltar que como indicador se tiene al número de años cumplidos Edad y la escala de medición de esta variable es cuantitativa y de tipo razón.

Tabla 5
Variable independiente Edad, dimensión e indicador

Variable	Dimensión	Indicador
Experiencia es la medida que aumenta el nivel de entrenamiento y especialización que hay en una empresa que aumentarán también las diferencias salariales entre el grueso de la población.	Años cumplidos	Número de años cumplidos

Otra variable independiente es la variable género, pero primero se debe aclarar antes de conceptualizarla, el enfoque que se consideró para incorporar esta variable en la presente investigación. Desde los informes sobre los salarios que realiza la Organización Internacional del Trabajo (OIT) desde el 2015, con una periodicidad bianual, se planteó que hay dos categorías para entender el análisis y la cuantificación de este concepto, la parte explicable y parte no explicable de una brecha salarial por género, brecha que muestra la diferencia salarial entre los ingresos que perciben las mujeres y los hombres, la primera señala el hecho de que las mujeres se insertan en el mercado laboral en forma diferente a los hombres respecto a variables observables tales como el nivel educativo alcanzado, categoría profesional, actividad económica y horas trabajadas, pero la otra parte, la no explicada es una que resulta de la resta del ajuste con esta características observables y que se constituye en la evidencia de la discriminación salarial relacionada con el hecho de ser mujer, por tanto se parte del entendido que cuando se contrasten los ingresos entre mujeres y hombres que realizan el mismo trabajo, se entiende en principio que en la actualidad debe existir una igualdad salarial cuando se realizan trabajos de igual valor (OIT, 2019).

Entiéndase que, si bien es cierto el concepto de género que está referido a las ideas, creencias y atribuciones sociales, que se van construyendo en cada cultura y momento de la historia de los pueblos, respecto a la diferencia sexual de sus integrantes, y que el concepto de sexo está referida a las diferencias y características biológicas, anatómicas, fisiológicas, entre otros, de los seres humanos, en esta investigación nos referimos como variable género, bajo el enfoque que en párrafo anterior se precisó, al hecho de ser mujer o hombre pero referido al concepto propuesto por la OIT “trabajo decente para todas las personas” (OIT, 2019).

A efectos de su operacionalización, se considera a esta variable como una variable dicotómica que se refiere al hecho de ser mujer o hombre, por tanto, la denominaremos como sexo.

La dimensión única de esta variable la denominaremos como sexo, en la que como indicador señalamos que será 0 si es mujer y será 1 si consideramos al hombre, y como escala de medición es una variable categórica de escala nominal.

Tabla 6
Variable independiente Género, dimensión e indicador

Variable	Dimensión	Indicador
La variable género está referida a lo propuesto por la OIT “trabajo decente para todas las personas”.	Sexo	0 = hombre 1 = mujer

Para una sociedad del conocimiento y de la información como en la que se vive actualmente, es fundamental contar con la preparación y capacitación de todos los profesionales en tecnologías de la información y comunicación, más conocidas como TICs, esto posibilita la comunicación, la interacción y las interconexiones entre los diferentes agentes económicos de una sociedad, en buena cuenta eleva su productividad al ejercer su actividad laboral, en ese sentido, es lógico pensar que al ser más productivo, es muy probable que tenga este hecho un efecto en la remuneración o en los salarios o en los ingresos en general que esta persona percibe, pero previo a revisar este impacto, es relevante precisar cómo se entiende las TICs en esta investigación, si sabemos que el uso de las TICs en las universidades del mundo son un elemento determinante en sus cambios y

adaptación a las nuevas formas de hacer y pensar en las diversas actividades en nuestras sociedades tal como lo plantea Ricardo y Iriarte (2017), o al hecho de que las universidades requieren formar profesionales que estén preparados para hacer uso de todo el ecosistema tecnológico que en la actualidad existe, en ese entender es que para la presente investigación se conceptúa a la variable TICs como “las herramientas para la gestión de la información que permiten y proporcionan soluciones y apoyo a la toma de decisiones” (Sieber, Valor, & Porta, 2006).

Para su operacionalización, para esta variable se considera si conoce o si utiliza o no, la computadora, en sus labores, esta forma de incorporar a la variable TICs, por ejemplo, entre otros estudios, la utilizó la investigación de Torres y Ochoa (2018) quienes en su investigación sobre desigualdad salarial asociada al uso de las tics realizaron para México.

Se cuenta con sólo una dimensión a la que se le denominó TICs, variable categórica binaria de escala nominal, con un indicador 0 = No y 1 = Si.

Tabla 7

Variable independiente TICs, dimensión e indicador

Variable	Dimensión	Indicador
La variable TICs se refiere a “las herramientas para la gestión de la información que permiten y proporcionan soluciones y apoyo a la toma de decisiones”	TICs	0 = No 1 = Si

3.3 Población, muestra, muestreo, unidad de análisis

En esta investigación la información que se utiliza para la evidencia empírica proviene de los datos generados por la encuesta nacional de hogares para el año 2019, que el Instituto Nacional de Estadística e Informática (INEI) realiza periódicamente desde 1995 con el propósito de efectuar el seguimiento de los indicadores sobre las condiciones de vida en el Perú y que tiene como cobertura a todo el país, tanto en el área urbana como en el rural, por tanto, población en la presente investigación es aquella que resulta de lo que nos permite inferir la encuesta del enaho, esto es, como los resultados tienen, en este caso, una periodicidad anual, estos tienen un alcance nacional, urbano, rural, departamental,

costa urbana, costa rural, sierra urbana, sierra rural, selva urbana, selva rural y a nivel del área metropolitana de Lima y Callao (INEI, 2022). Los criterios que se considera para incluir, son todas las viviendas particulares y sus ocupantes residentes en el área urbana y rural del país. Y los criterios de exclusión que se considera es a los miembros de las fuerzas armadas que viven en cuarteles, campamentos, barcos y otros, se excluye también a las personas que residen en viviendas colectivas como son los hoteles, hospitales, cárceles, asilos y claustros religiosos.

La muestra de este trabajo de investigación es el utilizado en la encuesta nacional de hogares (enaho), el tamaño de la muestra es de 128,276 observaciones que corresponden a 36,994 viviendas particulares, con 23,346 referidas al área urbana y 13,648 al área rural, tal como refiere los cálculos procesados en el stata y su ficha técnica (INEI, 2022).

Para el proceso de muestreo se consideró una muestra de tipo probabilística porque selecciona a todos los elementos de una población determinada al azar, de áreas porque considera las proporciones territoriales definidas geográficamente, estratificada por que se realiza una división previa a la población en estudio en grupos o clases homogéneas con respecto a alguna característica y se le asigna un peso, multietápica porque el proceso de selección es por etapas, primero es un centro poblado, luego un conglomerado y finalmente una vivienda, e independiente en cada departamento de estudio, y con un nivel de confianza de los resultados de 95% (INEI, 2022).

3.4 Técnicas e instrumentos de recolección de datos

Para la recolección de la información la metodología que utiliza la enaho es definir la unidad de investigación compuesta por los integrantes del hogar familiar, los trabajadores del hogar con cama adentro, los integrantes de una pensión familiar, y las personas que no son miembros del hogar familiar pero que estuvieron presentes el último mes, se encuesta al jefe de hogar, al cónyuge, a los perceptores de ingresos, y a las personas de 12 años a más, básicamente. Para la realización de esta encuesta utilizan como instrumento básico unas cartillas que son

cuestionarios que conforman 394 preguntas y 319 ítems organizadas de esta manera (INEI, 2022):

Tabla 8
Enaho, número de preguntas e ítems

	Preguntas	Ítems
Enaho. 01	142	
Enaho. 01-A	162	
Enaho. 01-B	43	Gastos 601 = 203
Enaho. 02	25	Gastos 602-612= 116
Enaho. 04	22	

Con el cuestionario elaborado y validado se conforman los equipos de trabajo de campo, con coordinadores departamentales, supervisores y encuestadores, quienes utilizando equipos móviles (tablets) capturan los datos a través de una entrevista directa, bajo el tipo de encuesta de derecho, quiere decir, que la población de estudio está constituida por todos los residentes habituales del hogar.

Cabe resaltar que con respecto a la validez y confiabilidad del instrumento, el cuestionario que se utiliza en esta encuesta, que mejor garantía que el análisis de la calidad de los resultados de la encuesta y de la metodología en general que utiliza la enaho, en base a un análisis del diseño y distribución de la muestra, de los resultados de las entrevistas, de las tasas de no respuesta total, no respuesta parcial, de la tasa de información referencial y de las correcciones a las no respuestas y de las correcciones a los valores extremos, explicadas en un documento publicado por el inei y que está referido al análisis de la calidad de la encuesta, en él están los resultados citados en la siguiente referencia (INEI, 2020).

3.5 Procedimiento

La información que brinda la encuesta enaho es amplia por tanto se ha procedido a explorar la data que permita procesar el trabajo de las variables relacionadas al tema de investigación como son, los ingresos que perciben las personas encuestadas, los niveles educativos alcanzados por estas personas, la edad, el sexo de las mismas, y si tienen o manejan la computadora, en ese sentido

presentamos una síntesis de los temas investigados en la encuesta enaho que va en la siguiente tabla elaborada en base al diccionario de la misma (INEI, 2020):

Tabla 9
Síntesis de los temas investigados por el enaho

No.	Temas	Preguntas
1.	Aspectos generales (Carátula)	(5 preguntas)
2.	Características de la vivienda y del hogar	(32 preguntas)
3.	Características de los miembros del hogar	(20 preguntas)
4.	Educación - Para personas de 3 años y más de edad	(46 preguntas)
5.	Salud - Para todas las personas	(29 preguntas)
6.	Empleo e Ingreso - Para personas de 14 años y más de edad	(63 preguntas)
7.	Sistema de Pensiones	(2 preguntas)
8.	Etnicidad	(3 preguntas)
9.	Desplazamiento de la población a otros distritos por trabajo	(1 pregunta)
10.	Inclusión Financiera	(4 preguntas)
11.	Gastos del hogar	(82 preguntas)
12.	Programas sociales de ayuda alimentaria	(7 preguntas)
13.	Programas sociales no alimentarios	(4 preguntas)
14.	Participación ciudadana	(6 preguntas)
15.	Módulo de opinión	(43 preguntas)

De ese total de temas investigados por esta encuesta, se identificó a las vinculadas al propósito de la presente investigación que están dentro de los temas 1, 3, 4 y 6.

Como la data de la encuesta está organizada por módulos (son 29 en total cuya lista va en el anexo 3), ahora seleccionamos aquellos que contienen estos temas y los presentamos de acuerdo a la siguiente tabla:

Tabla 10
Módulos del enaho a analizar

Código del módulo	Módulo
2	Características de los Miembros del Hogar
3	Educación
5	Empleo e Ingresos

La data en cada módulo también es amplia por tanto hemos seleccionado las pertinentes al propósito de la investigación, cabe señalar que en el módulo relacionado a las características de los miembros del hogar se tienen 20 preguntas, en el tema de la educación 46 preguntas, y en el tema relacionado al empleo y los ingresos existen 63 preguntas. Adicionalmente se considera una data relacionada al ubigeo, al dominio, estrato, y a los factores de expansión con los que se pueden

realizar el análisis inferencial, esto es llevar los resultados de la muestra al nivel de la población. Los resultados de este procedimiento son los que se muestran a continuación:

Tabla 11

Data del enaho a analizar por variables de la investigación

Variable	Enaho		
	Módulo	Variable	Etiqueta
1. Ingreso	5	P524A*	Ingreso total – monto en S/
		P529T*	Pago total (monto S/.) en especies dependiente
		P530A*	Ganancia neta en el mes anterior
		P536*	Valor de su consumo en el mes anterior
2. Educación	3	P301A	Último año o grado de estudios y nivel que aprobó
		P301B	
		P301C	
3. Edad	2	P208A	Edad en años cumplidos
4. Sexo	2	P207	Sexo
5. Tics	3	P314A	TICs
		P314B_2	
		P314B1_1	
		P314D	
		P316B	

* Si tuvo trabajo

Nota: Elaboración propia en base al Diccionario de datos de la encuesta enaho 2019.

Con la identificación de la data pertinente se procedió al análisis del mismo y para ello había que precisar el método seguido y la herramienta utilizada.

3.6 Método de análisis de datos

Para el procesamiento y análisis de la data especificada en la tabla 10 se utilizó la herramienta denominada Stata, en su versión Stata 16, software de procesamiento estadístico completo, y se siguió el procedimiento sugerido para el análisis de los datos, según la ruta cuantitativa de Hernández-Sampieri y Mendoza (2018).

Un primer aspecto que se consideró fue el de preparar la data por cada variable de acuerdo al requerimiento de nuestro modelo conceptual y al modelo econométrico, inicialmente en las ecuaciones (1) y (2), y posteriormente en la ecuación (3).

Con respecto a la variable dependiente ingreso, procedimos a calcular en base a la siguiente ecuación:

$$\text{Ing} = (\text{P524A1} + \text{P529T} + \text{P530A} + \text{P536}) / 12 \dots\dots \text{si P501} = 1 \quad (4)$$

Respecto a la variable educación, se tuvo que plantear ciertas condiciones a efectos de determinar el número de años dedicados a estudiar, y que se presenta en la siguiente ecuación:

$$\text{Educ} = \begin{cases} 0, & \text{si P301A} \leq 2 \\ \text{P301B}, & \text{si P301A} = 3 \text{ y P301B} \neq 0 \\ \text{P301C}, & \text{si P301A} = 3 \text{ y P301C} \neq . \\ 6, & \text{si P301A} = 4 \\ 6 + \text{P301B}, & \text{si P301A} = 5 \\ 11, & \text{si P301A} = 6 \\ 11 + \text{P301B}, & \text{si P301A} = 7 \\ 11 + \text{P301B}, & \text{si P301A} = 8 \\ 11 + \text{P301B}, & \text{si P301A} = 9 \\ 11 + \text{P301B}, & \text{si P301A} = 10 \\ 16 + \text{P301B}, & \text{si P301A} = 11 \end{cases} \quad (5)$$

En relación a la variable Edad, esta sería una variable cuantitativa discreta que la planteamos en relación a su data así:

$$\text{Edad} = \text{P208A} \quad (6)$$

En tanto que la variable Sexo, que vendría a ser una variable categórica nominal, la planteamos en la siguiente ecuación:

$$\text{Sexo} = \text{P207}, \text{ con la siguiente condición Mujer} = 0, \text{ Hombre} = 1 \quad (7)$$

Y, en relación a la variable Tics, la misma que se consideró para la segunda etapa de nuestro modelo (2) a efectos de medir su impacto en los ingresos, se considera en la ecuación (3) y se presenta como una variable categórica nominal en la siguiente ecuación:

$$\text{Tics} = \text{P316B}, \text{ con la siguiente condición } 0 = \text{No}, 1 = \text{Si} \quad (8)$$

El método econométrico (Wooldridge, 2010) que hemos adoptado es el de la regresión por mínimos cuadrados ordinarios a efectos de estimar los impactos de las variables independientes en la variable dependiente, el cambio en los ingresos ante cambios en las variables educación, edad, sexo, inicialmente y luego el cambio en los ingresos ante cambios de las variables referidas, más la variable

tics, para ello, se tuvo en cuenta que la relación entre una variable dependiente en términos de logaritmos frente a variables dependientes cuantitativas y categóricas, para su análisis y explicación se estiman sus coeficientes y por ejemplo el β_1 que vincula a la variación del ingreso con los años logrados de educación, darán como resultado, con el aumento en un año en educación, la variación en los ingresos aumentará en un porcentaje determinado.

Finalmente, cabe señalar, que lo primero que se ha hecho es una inspección de las variables del modelo antes de realizar la regresión, luego se efectuó la regresión, posteriormente se inspeccionó a los residuos a efectos de no tener problemas de heterocedasticidad, multicolinealidad o autocorrelación, los mismos que fueron tratados estadísticamente a efectos de tener resultados robustos.

3.7 Aspectos éticos

Hay lineamientos internacionales respecto a los criterios éticos a considerar en toda investigación científica, este trabajo de investigación se enmarca dentro de estos lineamientos. La Universidad César Vallejo a través de su Código de Ética en Investigación explica claramente la necesidad de mantener un marco deontológico en la investigación al igual que muchas instituciones tales como la declaración universal sobre bioética y derechos humanos de la Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura - UNESCO, la declaración en Singapur sobre la integridad en la investigación, la declaración de Helsinki, los aspectos éticos en la American British Educational Research Association – BERA, o la de la American Educational Research Association – AERA. (UCV, 2017)

Esta investigación hace suyo los siguientes principios:

Beneficencia: esta investigación espera beneficiar a los sujetos del estudio incluido el investigador

No maleficencia: esta investigación no generará daño o perjuicio, ni directa ni indirectamente a los participantes del estudio, ni en lo físico ni en lo psicológico

Autonomía: que todo participante de la presente investigación participe por su libre voluntad, sin restricciones ni impedimentos

Justicia: que se reciba un trato igualitario sin ningún tipo de exclusión

Nos sometemos al cumplimiento de estos principios éticos con absoluto respeto y claridad, conscientes de que en caso se produzca algún incumplimiento se asumirá las responsabilidades que disponga el Comité de Ética de la Universidad.

IV. RESULTADOS

A los resultados se los ha ordenado de la siguiente manera, primero se presentan las variables de la investigación construidas a partir de la encuesta enahho, luego las variables del modelo para esta primera etapa, en tercer lugar la exploración gráfica y estadística de las variables, luego los resultados del primer modelo, sus correcciones y su robustez y finalmente, los resultados del modelo en su segunda etapa, sus correcciones y su robustez.

En relación a la variable dependiente ingresos, los resultados obtenidos del procesamiento de la data del enahho dan como resultado a:

Tabla 12

Resultados del procesamiento de las variables de la investigación: Ing

Ingreso mensual	
Observaciones:	128,276
Rango:	[0, 0.26018.334]
Media:	999.245
Desviación estándar:	1270.1

Nota: síntesis de los resultados procesados en Stata que va en anexo 5.

En el que se observa que del total de la muestra el promedio de ingresos mensuales es de S/ 999.25 para el año 2019, sin embargo, hay un preocupante porcentaje de missing a evaluar.

En cuanto a las variables independientes, y específicamente en relación al nivel educativo logrado y luego del procesamiento de la data obtenida por la encuesta enahho en base a un criterio definido y explicado en la ecuación (5) de la presente investigación, se obtuvo el siguiente resultado:

Tabla 13

Resultados del procesamiento de las variables de la investigación: Educ

Educ - grado de estudios y nivel que aprobó	
Observaciones:	128,276
Rango:	[0, 18]
Media:	7.65103
Desviación estándar:	5.00579

Nota: síntesis de los resultados procesados en Stata que va en anexo 5.

Se observa que el nivel educativo promedio alcanzado por los encuestados en la muestra del estudio enaho, es de 7.6 años en promedio esto significa que la gran mayoría de los entrevistados en la muestra no llegan ni al nivel superior básico, están a mitad de secundaria la gran mayoría.

En cuanto a la variable Edad, los resultados se muestran en la siguiente tabla:

Tabla 14

Resultados del procesamiento de las variables de la investigación: Edad

Edad	
Observaciones:	128,276
Rango:	[0, 98]
Media:	33.7339
Desviación estándar:	22.8636

Nota: síntesis de los resultados procesados en Stata que va en anexo 5.

Se observa que el promedio de los entrevistados tiene 33.7 años esto significa que la gran mayoría es probable que cuente con experiencia laboral, por cuanto en el caso extremo de que se haya estudiado, lo que dice el promedio 8 años, a los 24 probablemente ya se estaba laborando y para los 34 registrados como promedio de edad en esta muestra, se podría asumir 11 años de experiencia, esto lo analizaremos en el siguiente capítulo.

Y, en relación a la variable sexo, los resultados son:

Tabla 15

Resultados del procesamiento de las variables de la investigación: Sexo

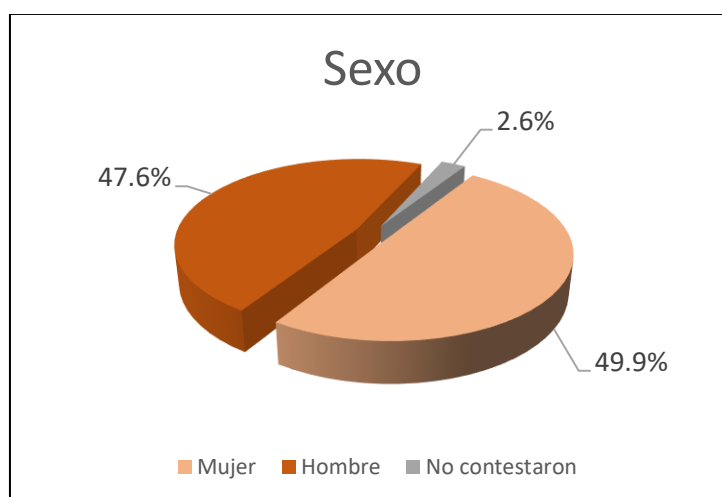
Sexo	
Observaciones:	128,276
Rango:	[0, 1]
Frecuencia	
Mujer (1):	63,954
Hombre (0):	61,025
No respondidos:	3,297

Nota: síntesis de los resultados procesados en Stata que va en anexo 5.

Como esta variable es una variable categórica, clasificada estadísticamente como una variable dummy, la presentamos gráficamente:

Figura 1

Resultados del procesamiento de las variables de la investigación: Sexo



Nota: síntesis de los resultados procesados en Stata que va en anexo 5.

En relación a las variables del modelo para la primera etapa de esta investigación sólo queda generar la variable logaritmo del ingreso y la variable edad al cuadrado, tal como hemos señalado en la ecuación (2) y tal como se señala en el marco teórico sobre la ecuación de Mincer. Los resultados para la variable $\log(\text{Ing})$ son:

Tabla 16

Creación de las variables de la investigación: Log(Ing)

Log(Ing) – logaritmo del ingreso mensual	
Observaciones:	128,276
Rango:	[0.5596158, 10.166556]
Media:	6.52062
Desviación estándar:	1.15236

Nota: síntesis de los resultados procesados en Stata que va en anexo 5.

Y, además la variable edad al cuadrado:

Tabla 17

Creación de las variables de la investigación: Edad²

Edad2 – edad al cuadrado	
Observaciones:	128,276
Rango:	[0, 9604]
Media:	1,660.72
Desviación estándar:	1,845.78

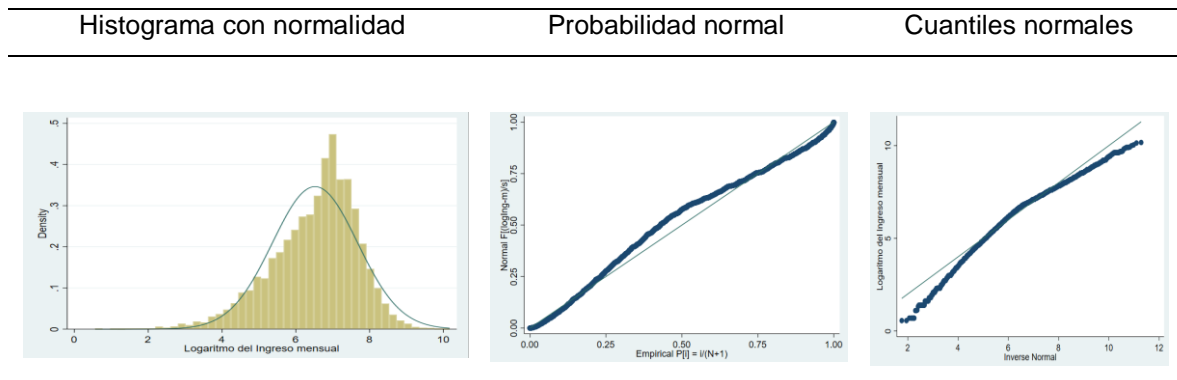
Nota: síntesis de los resultados procesados en Stata que va en anexo 5.

Ahora completadas todas las variables de la primera parte del modelo veremos los resultados de las pruebas de normalidad de cada variable, y de la asociación que hay entre la variable dependiente con cada variable independiente, antes de pasar a la estimación de la regresión.

Se utilizó dos conceptos, el análisis por visualización gráfica y el test de normalidad correspondiente a cada variable.

Con respecto a la variable dependiente ingreso, que en esta investigación está expresada en logaritmos, tiene un comportamiento similar a la distribución tal como se observa en los gráficos histograma y los de probabilidad normal y cuantiles normales:

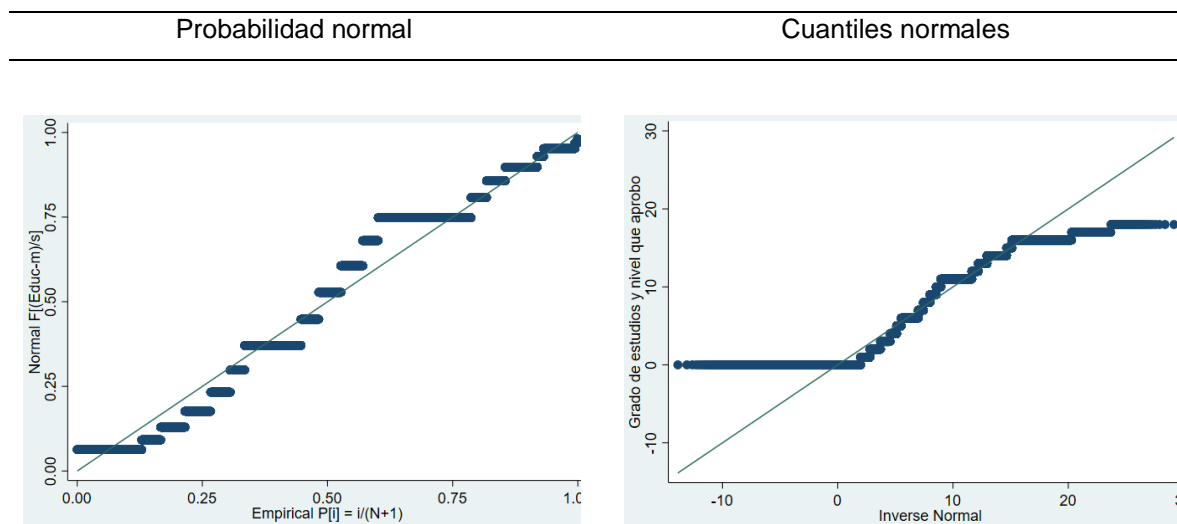
Figura 2
Pruebas gráficas de normalidad: $\log(\text{Ing})$



Nota: Elaboración y procesamiento propio en base a la encuesta de datos enaho 2019.

En tanto que para la variable educación se observa los siguientes resultados:

Figura 3
Pruebas gráficas de normalidad: *Educ*

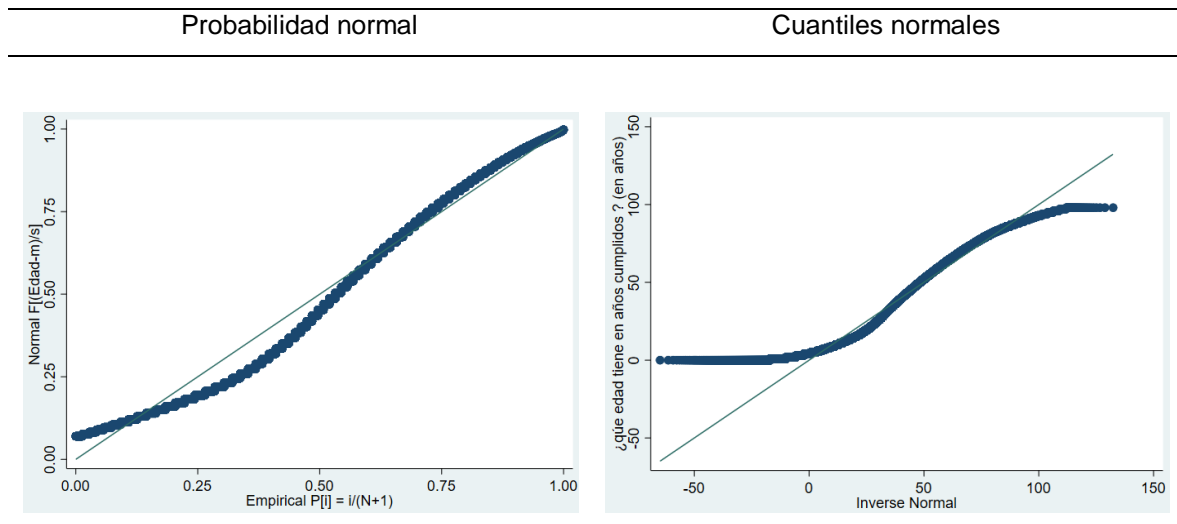


Nota: Elaboración y procesamiento propio en base a la encuesta de datos enaho 2019.

Para que los datos de la variable años de educación construidos a partir de la data de la encuesta sean normales, deberían estar pegados a la recta. Los gráficos de probabilidad normal no muestran indicios de no normalidad, mientras que los cuantiles normales muestran, una desviación de lo normal en las colas.

En cuanto a la variable experiencia expresada en la edad, se observa:

Figura 4
Pruebas gráficas de normalidad: Edad

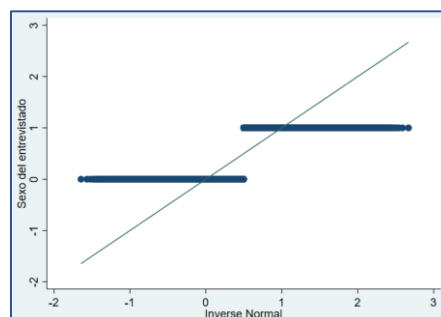


Nota: Elaboración y procesamiento propio en base a la encuesta de datos enaho 2019.

Los datos de la variable edad que representa la experiencia construidos a partir de la data de la encuesta sean normales, deberían estar pegados a la recta. Los gráficos de probabilidad normal no muestran indicios de no normalidad, y los cuantiles normales muestran, una desviación de lo normal en las colas.

Y, respecto a la variable sexo, como se trata de una variable categórica nominal se utiliza el gráfico de los cuantiles normales, y su lectura es diferente:

Figura 5
Pruebas gráficas de normalidad: Sexo*



- Es una variable categórica por tanto su análisis es diferente

Nota: Elaboración y procesamiento propio en base a la encuesta de datos enaho 2019.

En el gráfico se observa que la data se concentra en los dos valores asumidos para la identificación del sexo de los encuestados que en gráfico anterior se evidenció que el 49.9% eran mujeres en tanto que el 47.6% eran hombres.

Pero es pertinente en este tipo de variables utilizar las tablas cruzadas o de contingencia por que generan buena información y relevante para los propósitos de la investigación, es el caso si la cruzamos con los niveles educativos logrados, cuyos resultados se observan en la siguiente tabla:

Tabla 18
Tabla de contingencia: Sexo y Educ

Grado de estudios y nivel que aprobó	Sexo del entrevistado		Total
	Hombre	Mujer	
0	5,723	9,275	14,998
1	1,938	2,466	4,404
2	2,552	3,124	5,676
3	2,815	3,134	5,949
4	2,195	2,345	4,540
5	1,656	1,697	3,353
6	6,552	6,571	13,123
7	2,159	1,963	4,122
8	2,606	2,544	5,150
9	2,628	2,457	5,085
10	1,802	1,791	3,593
11	11,841	9,740	21,581
12	1,787	1,810	3,597
13	2,243	2,110	4,353
14	3,610	3,825	7,435
15	747	730	1,477
16	3,516	3,774	7,290
17	280	316	596
18	46	32	78
Total	56,696	59,704	116,400

Nota: síntesis de los resultados procesados en Stata que va en anexo 5.

Observamos que entre los encuestados el mayor porcentaje (18.5%) manifiestan haber culminado sus estudios secundarios especialmente los hombres (54.8%), el 12.9% manifiestan no haber cursado ningún estudio especialmente las mujeres (61.8%) y que el 11.3% manifiestan haber culminado sus estudios de primaria.

Inspeccionando las variables del modelo antes de la regresión obtuvimos los siguientes resultados de su estadística básica:

Tabla 19
Estadísticas básicas de las variables del modelo (2)

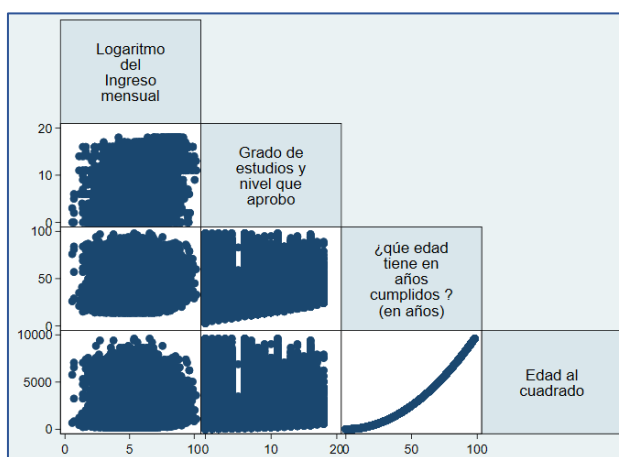
Variable	Obs	Media	Desv.estandar	Min	Max
logIng	56,101	6.520618	1.152358	.5596158	10.16656
Educ	116,400	7.651031	5.005789	0	18
Edad	124,979	33.73394	22.86356	0	98
Edad2	124,979	1660.717	1845.779	0	9604
Sexo	124,979	.511718	.4998647	0	1

Nota: síntesis de los resultados procesados en Stata que va en anexo 5.

Se observan por variable el número de observaciones, los promedios, la desviación estándar, y los valores mínimo y máximo.

Y también se graficó la matriz de dispersión a efectos de ver el comportamiento de cada variable y se obtuvo el siguiente resultado:

Figura 6
Gráfico de la matriz de dispersión



Nota: Elaboración y procesamiento propio en base a la encuesta de datos enaho 2019.

Así como la matriz de correlación a efectos de ver la asociación existente entre la variable dependiente y cada una de las variables independientes, obtuvo el siguiente resultado:

Tabla 20

Matriz de correlación entre variable dependiente y variables independientes

	logIng	Educ	Edad	Edad2
logIng	1.0000			
Educ	0.4541* 0.0000	1.0000		
Edad	-0.1147* 0.0000	0.1086* 0.0000	1.0000	
Edad2	-0.1596* 0.0000	-0.0383* 0.0000	0.9590* 0.0000	1.0000

Nota: procesamiento propio en base a la encuesta de datos enaho 2019, resultado en anexo 5.

Los coeficientes de correlación significativos al 95% ((1-0.05)*100%) de confianza se muestran con una estrella*, los resultados nos señalan que hay correlación entre la variable dependiente logIng y la variable Educ, su pvalor es menor al 0.05, de igual manera para la correlación entre la variable logIng y la variable Edad y la variable logIng y la variable Edad2.

Ahora efectuamos la regresión por mínimos cuadrados ordinarios de nuestro modelo planteado en la ecuación (2):

$$\log \text{Ing} = \beta_0 + \beta_1 \text{Educ} + \beta_2 \text{Edad} + \beta_3 \text{Edad}^2 + \beta_4 \text{Sexo} + \varepsilon$$

Los resultados son los que se muestran en la tabla 20, en ella se observa, que la varianza total tiene n-1 grados de libertad, en este caso 56,085 observaciones, los predictores del modelo son las cuatro variables independientes más una constante, por tanto es 4, pero a efectos de evaluar la relación estadísticamente significativa entre las variables independientes y la variable dependiente, el pvalor obtenido en esta regresión debe ser menor a 0.05 y eso se revisa en **Prob>F**, y como observamos en los resultados efectivamente el pvalor es una cifra muy pequeña y menor a 0.05, lo que nos indica la significancia conjunta de las variables (Educ, Edad, Edad2 y Sexo) para explicar la variable dependiente log(Ing).

Tabla 21
Resultados de la regresión del modelo (2)

Source	SS	df	MS	Obs	=	56,086
Model	20104.6942	4	5026.174	F (4, 56081)	=	5183.77
Residual	54378.0595	56081	0.969599	Prob > F	=	0.0000
				R - squared	=	0.2699
				Adj R - squared	=	0.2699
Total	74482.7537	56085	1.327998	Root MSE	=	.98468

logIng	Coef	Std. Err	t	P > τ	[95% conf.	Interval]
Educ	.1108613	.0009929	111.65	0.000	.1089152	.1128075
Edad	.0701489	.0013296	52.76	0.000	.0675429	.072755
Edad2	-.0007134	.0000142	-50.19	0.000	-.0007413	-.0006856
Sexo	-.397677	.0084339	-47.15	0.000	-.4142074	-.3811465
_cons	4.118616	.0309389	133.12	0.000	4.057975	4.179256

Nota: procesamiento propio en base a la encuesta de datos enah 2019, resultado en anexo 5.

Otro indicador clave es el R-cuadrado ajustado que en este caso mide lo mismo que el R-cuadrado, cuando el número de variables es pequeño como en este caso y el número de observaciones es muy grande como en este caso, entonces el R-cuadrado ajustado está más cerca que el R-cuadrado, eso señala una asociación más honesta entre las variables independientes y la variable dependiente.

El otro indicador clave es el **Root MSE** que es la desviación estándar del error, y es la raíz cuadrada del MS Residual, que cuanto más cercano a cero será mejor el ajuste del modelo, en este caso este indicador no es favorable al modelo.

Respecto a los pvalor para cada coeficiente del modelo son significativos todos puesto que sus resultados señalan ser menores a 0.05, esto indica que cada una de las variables son estadísticamente significativas para explicar el log(Ing).

Con las atingencias planteadas, los resultados del modelo indican para la variable Educ, que si se aumenta en un año la educación los ingresos aumentan en 11%.

Para la variable Edad, que representa a la experiencia, los resultados nos indican a mayor experiencia los ingresos se incrementan en 7%.

Pero para la variable Sexo, los resultados señalan que por ser mujer sus ingresos son menores en 39.8% respecto a lo que perciben los hombres.

Finalmente, se analizó que se cumplan los supuestos más importantes de la regresión (Wooldridge, 2010), que no tenga problemas de heterocedasticidad, multicolinealidad y que los residuos se comporten normalmente. Para la heterocedasticidad se aplica un test cuya hipótesis nula es que existe homocedasticidad en los residuos, el test es de Breusch-Pagan y los resultados son:

Tabla 22
Test de Breusch – Pagan

Ho:	Existe homocedasticidad
Variables:	Valores de ajuste del log(Ing)
Chi2(1)	914.52
Prob > chi2	0.0000

Nota: procesamiento propio en base a la encuesta de datos enaho 2019, resultado en anexo 5.

La hipótesis nula dice que existe homocedasticidad en los residuos, los resultados muestran que se rechaza la hipótesis nula porque la probabilidad es menor a 0.05 y eso significa que hay heterocedasticidad o dispersión desigual de los residuos porque la regresión de mínimos cuadrados ordinarios supone que todos los residuos se extraen de una población que tiene varianza constante, eso se llama homocedasticidad. Esto se controlará mejorando la regresión con el concepto de robustez en las regresiones.

Otro supuesto importante para el modelo de regresión múltiple es que las variables independientes no deben estar correlacionadas entre sí, una no debería ser función lineal de la otra, si esto sucede los errores estándar se pueden incrementar, para analizar este supuesto hay una prueba que se denomina factor de inflación de la varianza (vif), si una variable tiene $vif > 10$ o $1/vif < 0.1$ es señal de multicolinealidad (Wooldridge, 2010), los resultados muestran que:

Tabla 23

Test de factor de inflación de la varianza (vif)

Variable	VIF	1/VIF
Edad2	26.99	0.037052
Edad	26.46	0.037797
Educ	1.21	0.826676
Sexo	1.00	0.996356
Media VIF	13.91	

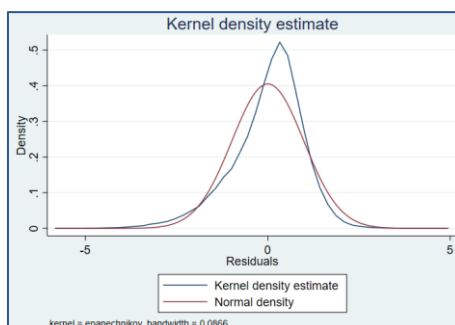
Nota: procesamiento propio en base a la encuesta de datos enaho 2019, resultado en anexo 5.

Observamos que tanto la variable edad2 y edad tienen un vif mayor a 10, es porque la variable edad2 es el cuadrado de la variable edad, pero esto es resultado de la propuesta teórica en la ecuación de Mincer en el que se señala que a medida que avanzan los años se produce un efecto que es el de los rendimientos decrecientes y que se penaliza en los ingresos, dicho esto cabe indicar que las otras variables no presentan el problema de multicolinealidad.

Y el otro supuesto en el modelo de regresión es que los residuos se distribuyen normalmente, los residuos son la diferencia entre los valores observados y los predichos, y para comprobar ello utilizamos la prueba de densidad de Kernel (Wooldridge, 2010), cuyo resultado se presenta a continuación:

Figura 7

Gráfico de Kernel, normalidad de los residuos



Nota: Elaboración y procesamiento propio en base a la encuesta de datos enaho 2019.

Para verificar que el comportamiento de una variable es similar al de una distribución normal en este caso es conveniente utilizar los gráficos sobre los test por la cantidad de observaciones utilizadas en esta investigación, en ese sentido,

observamos en la Figura 7 que la distribución de los residuos se asemeja a una distribución normal.

Finalmente, a efectos de controlar la heterocedasticidad y que el modelo sea estadísticamente robusto, volvemos a estimar la regresión incluyendo este concepto que permite que los errores estándar robustos afecten positivamente a los valores t e intervalos de confianza y mejore la significancia estadística del modelo.

Tabla 24
Regresión robusta del modelo (2)

Regresión lineal		Obs	=	56,086		
		F (4, 56081)	=	4860.6		
		Prob > F	=	0.0000		
		R - squared	=	0.2699		
		Root MSE	=	.98468		
Robust						
logIng	Coef	Std. Err	t	P > τ	[95% conf.	Interval]
Educ	.1108613	.0010065	110.15	0.000	.1088887	.112834
Edad	.0701489	.0014518	48.32	0.000	.0673034	.0729945
Edad2	-.0007134	.0000158	-45.03	0.000	-.0007445	-.0006824
Sexo	-.397677	.0086205	-46.13	0.000	-.4145733	-.3807806
_cons	4.118616	.0330355	124.67	0.000	4.053866	4.183366

Nota: procesamiento propio en base a la encuesta de datos enah 2019, resultado en anexo 5.

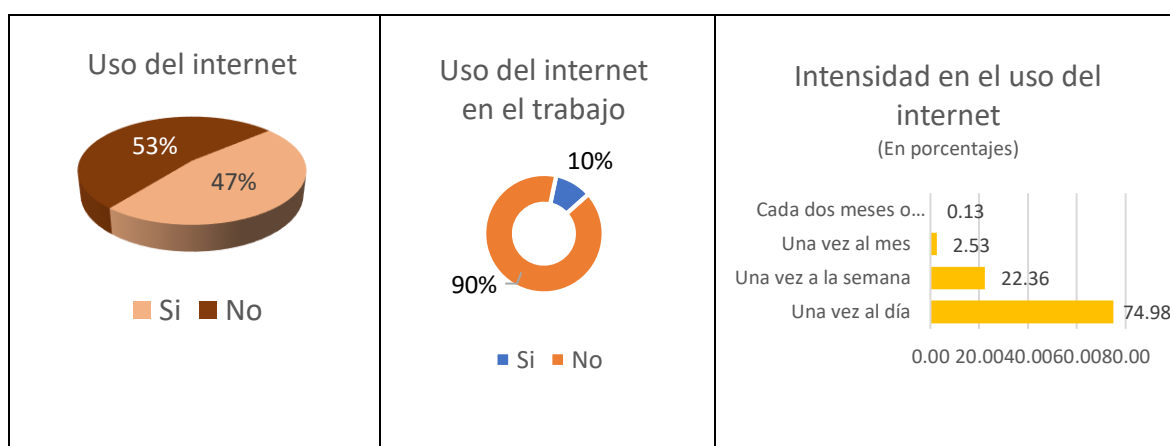
Los resultados muestran que se mantienen los valores de los coeficientes, la prob > F es significativa, las pruebas t han sido mejoradas en todas las variables y los intervalos se han ajustado, con lo que se puede afirmar que esta regresión es estadísticamente significativa y sus resultados son estadísticamente robustos.

En la segunda etapa del modelo, vamos a considerar una nueva variable independiente, es la referida a las tecnologías de comunicación e información, que se la ha denominado TICs, y para construir esta variable a partir de los datos

proporcionados por la encuesta enahoy, es pertinente presentar la siguiente información como representativa del uso de las TICs en el Perú, los datos relacionados a el uso del internet y el uso de la computadora.

Con respecto al internet los resultados son los que se observa en la siguiente figura:

Figura 8
Resultados del uso del internet



Nota: Elaboración y procesamiento propio en base a la encuesta de datos enahoy 2019.

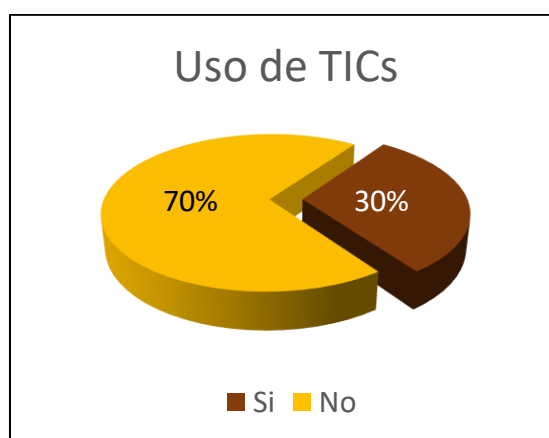
Se observa que entre los encuestados las respuestas que brindan señalan que hay un menor uso del internet, solo el 47% afirma haber usado el internet en los últimos 30 días, en tanto que el 53% manifiesta que no utilizaron. En relación al uso del internet en el trabajo los encuestados que respondieron esta pregunta, sólo el 10% de ellos manifiestan que, si usaron el internet en el trabajo y el 90%, manifestaron que no han utilizado el internet en el trabajo, los que respondieron que si han utilizado el internet en los últimos 30 días también señalaron que lo utilizan una vez al día (el 75%), o una vez a la semana (el 22.4%).

Cabe mencionar que, en estas preguntas relacionadas a los temas del internet, se presenta en los resultados de la encuesta del enahoy, un alto porcentaje de encuestados que no han respondido y por tanto el software stata los considera como valores perdidos (missing), sin embargo, esa presencia de missing no invalida los resultados hallados y tampoco afecta a su significancia estadística.

En esta investigación para construir la variable TICs a partir de los datos proporcionados por el enaho 2019, se considera a la pregunta P316B que dice si en los últimos 3 meses ¿ha utilizado una computadora, laptop, tablet o similar? (INEI, 2022) por tanto esta variable ingresa a nuestro modelo econométrico (3) como una variable categórica binaria nominal, con dos posibilidades que haya utilizado la computadora, la respuesta sería 1 = Si, o que no haya utilizado y la respuesta sería 0 = No, los resultados obtenidos son los que se presentan en la siguiente figura:

Figura 9

Resultados del procesamiento de las variables de la investigación: TICs



Nota: Elaboración y procesamiento propio en base a la encuesta de datos enaho 2019.

Como se puede observar sólo el 30% de los encuestados que han respondido, manifiestan que, si utilizan la computadora, para un mayor análisis y como esta variable es categórica, utilizamos a la tabla de contingencia o cruzada en primer lugar con el indicador sexo y luego con la variable educación y este es el resultado:

Tabla 25

Tabla de contingencia: TICs y Sexo

Uso de TICs	Sexo del entrevistado		Total
	Hombre	Mujer	
No	29,719	34,363	64,082
Si	14,635	13,344	27,979
Total	44,354	47,707	92,061

Nota: procesamiento propio en base a la encuesta de datos enaho 2019, resultado en anexo 5.

De las personas que han respondido que no utilizan la computadora el 54% son mujeres y 46% son hombres, y de las personas que han respondido que si utilizan la computadora el 52% son hombres y el 48% mujeres. Entre los hombres que han contestado a la encuesta el 67% manifiestan no usar la computadora y el 33% si utilizan la computadora, mientras que entre las mujeres que han contestado la encuesta el 72% dicen no utilizar la computadora sólo el 28% manifiestan que si utilizan la computadora.

La otra tabla de contingencia que se ha procesado es la que relaciona a la educación lograda con el uso de las TICs, y este es el resultado:

Tabla 26
Tabla de contingencia: Educ y TICs

Grado de estudios y nivel que aprobó	Uso de TICs		Total
	No	Si	
0	6,433	9	6,442
1	2,106	5	2,111
2	3,129	21	3,150
3	3,403	28	3,431
4	2,134	32	2,166
5	1,063	37	1,100
6	10,758	314	11,072
7	2,115	505	2,620
8	3,342	1,443	4,785
9	3,502	1,572	5,074
10	1,936	1,657	3,593
11	16,315	5,266	21,581
12	1,297	2,300	3,597
13	1,191	3,162	4,353
14	3,143	4,292	7,435
15	301	1,176	1,477
16	1,732	5,558	7,290
17	75	521	596
18	6	72	78
Total	63,981	27,970	91,951

Nota: procesamiento propio en base a la encuesta de datos enah 2019, resultado en anexo 5.

De solo el 30% de todos los encuestados que manifiestan utilizar la computadora, se observa que a medida que se tiene un mayor logro educativo, se utiliza más la computadora, en especial en el nivel universitario los que el 80% manejan la computadora, dato relevante para el propósito de la investigación por que ahora analizaremos el impacto que este hecho significa en los ingresos de estos universitarios.

Planteamos nuestro modelo econométrico (3) en la siguiente ecuación:

$$\log \text{Ing} = \beta_0 + \beta_1 \text{Educ} + \beta_2 \text{Edad} + \beta_3 \text{Edad}^2 + \beta_4 \text{Sexo} + \beta_5 \text{TICs} + \varepsilon$$

Presentamos las estadísticas básicas de las variables del modelo (3):

Tabla 27

Estadísticas básicas de las variables del modelo (3)

Variable	Obs	Media	Desv Están	Min	Max
logIng	56,101	6.520618	1.152358	.5596158	10.16656
Educ	116,400	7.651031	5.005789	0	18
Edad	124,979	33.73394	22.86356	0	98
Edad2	124,979	1660.717	1845.779	0	9604
Sexo	124,979	.511718	.4998647	0	1
TICs	92,061	.3039181	.4599502	0	1

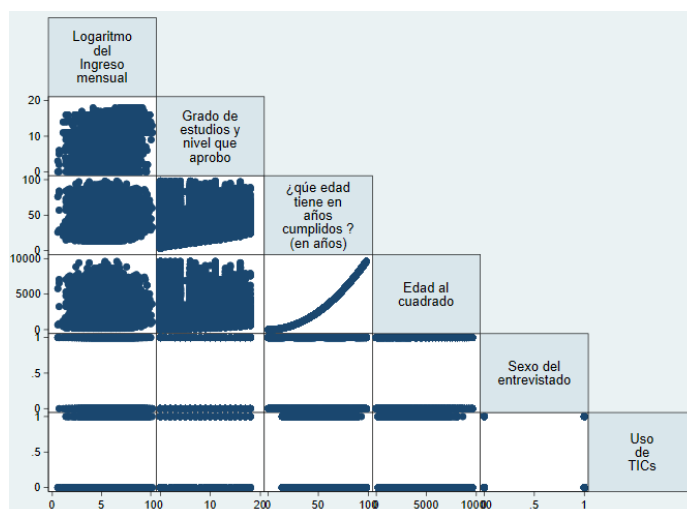
Nota: procesamiento propio en base a la encuesta de datos enah 2019, resultado en anexo 5.

Se observa por cada variable el número de observaciones, los promedios, la desviación estándar, y los valores máximo y mínimo.

Se obtuvo la matriz de dispersión y el gráfico correspondiente a esta matriz de las variables del modelo (3) con el siguiente resultado:

Figura 10

Gráfico de la matriz de dispersión de las variables del modelo (3)



Nota: procesamiento propio en base a la encuesta de datos enaho 2019, resultado en anexo 5.

También se obtuvo la matriz de correlación de las variables del modelo (3) para observar la asociación existente entre la variable dependiente y cada una de las variables independientes, el resultado es el siguiente:

Tabla 28

Matriz de correlación entre variable dependiente y variables independientes modelo (3)

	logIng	Educ	Edad	Edad2	Sexo	TICs
logIng	1.0000					
Educ	0.4541*	1.0000				
Edad	-0.1147*	0.1086*	1.0000			
Edad2	-0.1596*	-0.0383*	0.9590*	1.0000		
Sexo	-0.1814*	-0.0692*	0.0294*	0.0262*	1.0000	
TICs	0.3271*	0.5096*	-0.3773*	-0.3474*	-0.0546*	1.0000

Nota: procesamiento propio en base a la encuesta de datos enaho 2019, resultado en anexo 5.

Estos resultados muestran que hay correlación entre la variable dependiente $\log(\text{Ing})$ y cada una de las variables independientes, toda vez que sus p valor son menores al 0.05.

Se procedió a estimar la regresión por mínimos cuadrados ordinarios el modelo econométrico (3) y estos son los resultados:

Tabla 29
Resultados de la regresión del modelo (3)

Source	SS	df	MS	Obs	=	56,086
				F (5, 56080)	=	4496.74
Model	21315.2175	5	4263.04349	Prob > F	=	0.0000
Residual	53165.5362	56,080	.948030246	R - squared	=	0.2862
				Adj R - squared	=	0.2861
Total	74480.7537	56,085	1.32799775	Root MSE	=	.97367

logIng	Coef	Std. Err	t	P > τ	[95% conf.	Interval]
Educ	.0894523	.0011502	77.77	0.000	.087198	.0917067
Edad	.0774528	.0013305	58.21	0.000	.0748449	.0800606
Edad2	-.0007834	.0000142	-55.21	0.000	-.0008112	-.0007556
Sexo	-.4111998	.0083481	-49.26	0.000	-.4275621	-.3948374
TICs	.3959572	.0110808	35.73	0.000	.3742387	.4176758
_cons	4.043689	.0306647	131.87	0.000	3.983586	4.103792

Nota: procesamiento propio en base a la encuesta de datos enano 2019, resultado en anexo 5.

Los resultados del modelo (3) presentan una significancia conjunta de las variables Educ, Edad, Edad2, Sexo y TICs para explicar la variable dependiente observada en el valor $\text{Prob}>F$. El R-cuadrado ajustado tiene un valor muy similar al R-cuadrado, pero el valor root mse no es favorable al modelo.

Todos los p valor de los coeficientes se muestran significativos toda vez que son menores a 0.05, esto indica que cada una de las variables son estadísticamente significativas para explicar la variable dependiente.

Los resultados del modelo (3) señalan lo siguiente, si aumenta en un año la educación el impacto en los ingresos es el aumento en 9%, a más experiencia esto impacta en los ingresos en 8%, por ser mujer los ingresos son menores en 41%

respecto a lo que perciben los hombres y si utilizan TICs esto impacta positivamente en sus ingresos, en 40%, que si no lo utilizan.

Tal como se analizó el modelo anterior respecto a sus principales supuestos, también se analizó este modelo (3) (Wooldridge, 2010), lo importante es controlar el problema de heterocedasticidad, si tuviese algún problema de multicolinealidad y que los residuos se distribuyan normalmente.

Se realizó el test de Breusch – Pagan y este fue el resultado:

Tabla 30
Test de Breusch – Pagan modelo (3)

Ho:	Existe homocedasticidad
Variables:	Valores de ajuste del log(Ing)
Chi2(1)	1218.58
Prob > chi2	0.0000

Nota: procesamiento propio en base a la encuesta de datos enah 2019, resultado en anexo 5.

Los resultados señalan que tiene problemas de heterocedasticidad y para controlar este problema utilizaremos el concepto de robustez.

También se analizó la presencia de multicolinealidad a través de la prueba del factor de inflación de la varianza (vif) y estos son los resultados:

Tabla 31
Test de factor de inflación de la varianza (vif) del modelo (3)

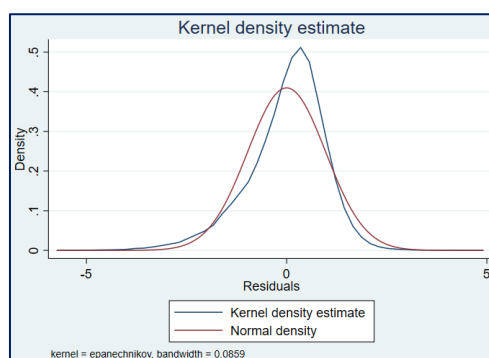
Variable	VIF	1/VIF
Edad2	27.51	0.036346
Edad	27.10	0.036905
Educ	1.66	0.602367
Sexo	1.01	0.994308
TICs	1.50	0.667870
Media VIF	11.75	

Nota: procesamiento propio en base a la encuesta de datos enah 2019, resultado en anexo 5.

Los resultados indican que salvo las variables Educ y Educ2 que conceptualmente están relacionadas, las otras variables no presentan evidencia de multicolinealidad.

Para el supuesto de que los residuos se distribuyen normalmente se utilizó la prueba de densidad de Kernel (Wooldridge, 2010), y este es el resultado:

Figura 11
Gráfico de Kernel, normalidad de los residuos del modelo (3)



Nota: procesamiento propio en base a la encuesta de datos enah0 2019, resultado en anexo 5.

Se observa que la distribución de los residuos se asemeja a una distribución normal.

Finalmente, volvemos a estimar la regresión del modelo (3) pero incluyendo el concepto de robustez, y estos son los resultados que señalan claramente, que el modelo en su conjunto es estadísticamente significativo, que todas las pruebas t han sido mejoradas en todas las variables controlándose el tema de la heterocedasticidad, que no tiene problemas de multicolinealidad y que los residuos se asemejan al de una distribución normal, y además comparativamente este modelo tiene un mejor ajuste que el modelo (2) y esto se visualiza en el R-cuadrado ajustado.

Tabla 32
Regresión robusta del modelo (3)

Regresión lineal		Obs	=	56,086		
		F (5, 56080)	=	4428.64		
		Prob > F	=	0.0000		
		R - squared	=	0.2862		
		Root MSE	=	.97367		
logIng	Robust					
	Coef	Std. Err	t	P > t	[95% conf.	Interval]
Educ	.0894523	.0011848	75.50	0.000	.0871302	.0917745
Edad	.0774528	.0014658	52.84	0.000	.0745799	.0803257
Edad2	-.0007834	.0000159	-49.20	0.000	-.0008146	-.0007522
Sexo	-.4111998	.0085365	-48.17	0.000	-.4279314	-.3944681
TICs	.3959572	.0107496	36.83	0.000	.374888	.4170265
_cons	4.043689	.0332553	121.60	0.000	3.978509	4.10887

Nota: procesamiento propio en base a la encuesta de datos enaho 2019, resultado en anexo 5.

V. DISCUSIÓN

El objetivo en la presente investigación es demostrar que la hipótesis general que afirma que “existe un impacto positivo, del nivel educativo logrado y la experiencia de los egresados universitarios, en el nivel de ingresos de estos en su actividad laboral, y un impacto diferenciado entre el género de los egresados universitarios con el nivel de ingresos de estos en su actividad laboral”, en esta parte de la investigación lo que desarrollamos es la contrastación entre lo propuesto conceptualmente y los resultados obtenidos en la investigación cuantitativa a partir de la data proporcionada por la encuesta nacional de hogares enaho para el Perú.

Para ello el modelo que se ha planteado en esta investigación corresponde a la ecuación de Mincer, adecuado a los datos que nos ofrece la encuesta enaho, cuyas principales variables cuantitativas son la variable dependiente ingresos, y las variables independientes educación y edad (como variable proxy de la experiencia), y la variable independiente sexo considerada como variable dummy, por ser categórica.

Los principales hallazgos a partir de una información procesada y estadísticamente significativa y robusta, demostrada en el capítulo anterior, son los siguientes:

Tabla 33

Estadísticas básicas de las variables cuantitativas del modelo (2)

Estadísticos	Ing	Educ	Edad
N	64,954	116400	124979
Max	26,018.33	18	98
Mín	0	0	0
Rango	26,018.33	18	98
Media	999.2453	7.651031	33.73394
Des Están.	1,270.104	5.005789	22.86356
CV	1.271063	.6542633	.6777613

Nota: procesamiento propio en base a la encuesta de datos enaho 2019, resultado en anexo 5.

En él se obtiene que el promedio de ingresos mensuales en el 2019 en el Perú, de acuerdo a la data proporcionada por la encuesta enahop es de S/999.24 soles mensuales.

Que el promedio general de edad de los encuestados es de 33.7 años por los que se puede asumir que se cuenta con experiencia laboral, que el promedio general del nivel educativo en el Perú, para el 2019, era de 7.6 años de escolaridad de los encuestados en la muestra de la encuesta enahop, esto significaría que mayormente la población apenas obtiene niveles educativos del inicio de la estudios de la secundaria, para profundizar aún más este resultado se obtuvo una tabla cruzada con la variable categórica sexo y se obtuvo información más esclarecedora en la tabla 17, y para efectos de un análisis más acucioso, procedimos a utilizar el factor de expansión para determinar su alcance a nivel de la población a nivel nacional, y esto se obtuvo:

Tabla 34
Tabla cruzada de Educ y Sexo a nivel población

Grado de estudios y nivel que aprobó	Sexo del entrevistado		Total
	Hombre	Mujer	
0	238,703.5	780,592.0	1,019,295.5
1	128,991.6	253,880.6	382,872.2
2	229,639.4	348,031.5	577,670.9
3	305,993.6	388,039.9	694,033.5
4	194,839.7	238,411.1	433,250.8
5	118,700.5	137,376.6	256,077.1
6	1,179,824.1	1,249,111.9	2,428,936.0
7	347,865.3	304,166.86	652,032.2
8	667,723.4	641,780.3	1,309,503.7
9	745,145.5	679,573.91	1,424,719.4
10	534,019.7	529,786.8	1,063,806.5
11	3,804,250.3	3,015,702.9	6,819,953.2
12	582,745.5	584,679.94	1,167,425.4
13	707,826.6	667,359.61	1,375,186.2
14	1,215,290.1	1,233,566.9	2,448,857.0
15	238,464.7	233,318.9	471,783.6
16	1,053,819.1	1,071,562.8	2,125,381.9
17	78,891.4	80,687.13	159,578.5
18	15,338.3	10,507.51	25,845.8
Total	12,388,072	12,448,137	24,836,209

Nota: procesamiento propio en base a la encuesta de datos enahop 2019, resultado en anexo 5.

Y con esta tabla obtenemos información muy valiosa para la investigación, como la población del estudio que sería 24.8 millones de habitantes, de los cuales 1.2 millones estarían en el nivel universitario, sin embargo, con la data proporcionada por sunedu, sólo en el sistema universitario, para el 2019, figuran 1.3 millones de alumnos universitarios, y que la gran mayoría de la población en el Perú, esto es 6.8 millones, estos han logrado concluir solamente sus estudios de secundaria.

Luego se procedió a estimar a través de una regresión con el método de los mínimos cuadrados ordinarios y con un modelo estadísticamente significativo y robusto obteniéndose los resultados que los contrastamos con la hipótesis general de la investigación:

Tabla 35
Resultados de la regresión del modelo (2)

Resultados:	
Variable	Coficiente
Educ	0.1108613
Edad	0.0701489
Edad2	-0.0007134
Sexo	-0.397677

Nota: procesamiento propio en base a la encuesta de datos enaho 2019, resultado en anexo 5.

Estos resultados contrastados con la hipótesis general de la investigación prueban que las variables educación y edad (proxy de experiencia) tienen un impacto positivo en la variable dependiente ingreso, significando además que con un año de más estudios se obtendría 11% más en sus ingresos, de igual manera sucede con la variable edad, variable proxy de experiencia, tiene un impacto positivo en los ingresos, esto es con más experiencia se obtiene un ingreso diferenciado, el impacto es de 7% más en los ingresos.

Los resultados obtenidos en esta regresión por la variable sexo confirman la propuesta de hipótesis de la investigación en que se obtendría un impacto diferenciado por ser mujer u hombre, los resultados lo que evidencian son que por ser mujer se tiene un ingreso menor en 39.8% respecto a los que obtiene un hombre como ingresos en el Perú.

Con respecto a los resultados obtenidos en la segunda etapa del modelo conceptual de la investigación, esto es, el modelo econométrico (3), después de todos los análisis de los supuestos y garantizada la significancia estadística y su robustez, en esta parte contrastaremos la hipótesis con los resultados y compararemos los dos modelos robustos hallados.

Sobre lo primero, la hipótesis propuesta señala que “existe un impacto positivo, del nivel educativo logrado, el manejo de las TICs y la experiencia, de los egresados universitarios, en el nivel de ingresos de estos en su actividad laboral, y un impacto diferenciado entre el género de los egresados universitarios con el nivel de ingresos de estos en su actividad laboral”, para confirmar esta hipótesis dos aspectos a analizar, una referida a los hallazgos obtenidos en el procesamiento de la información referida a las TICs, y el otro referido a los resultados de los coeficientes obtenidos en la estimación de la regresión del modelo (3), con respecto al primero, cuyos resultados los presentamos en la figura 9, tabla 25, y la siguiente tabla claramente se obtuvo que a mayores niveles educativos logrados más uso de las TICs, en el nivel universitario el 80% utiliza las TICs

Tabla 36
Resultados de la tabla resumen de contingencia: Educ y TICs del modelo (3)

Grado de estudios y nivel que aprobó	Uso de TICs	
11	5266	
12	2300	
13	3162	
14	4292	79.9%
15	1176	
16	5558	
17	521	
18	72	
Nivel universitario	22,347	
Muestra total	27,970	

Nota: procesamiento propio en base a la encuesta de datos enahó 2019, resultado en anexo 5.

Con relación a lo propuesto por la hipótesis y los resultados de los coeficientes de la estimación de la regresión del modelo (3), los coeficientes han tenido variaciones en cuanto al valor de los coeficientes del modelo anterior, pero

no respecto al sentido de la causalidad, ello confirma la hipótesis de la investigación para esta segunda etapa, los resultados los presentamos en la siguiente tabla resumen que va a continuación:

Tabla 37
Resultados de la regresión del modelo (3)

Resultados:	
Variable	Coficiente
Educ	0.0894523
Edad	0.0774528
Edad2	-0.0007834
Sexo	-0.4111998
TICs	0.3959572

Nota: procesamiento propio en base a la encuesta de datos enahó 2019, resultado en anexo 5.

Estos resultados confirman la relación positiva entre la variable ingresos y las variables educación, experiencia, y uso de TICs, con sus respectivos impactos, un año más de educación genera un diferencial de ingresos mayor en 9%, más experiencia genera un impacto en sus ingresos de 8% más, el uso de las TICs genera un impacto positivo en sus ingresos en 40%. También confirma que existe un impacto diferencial entre ingresos por concepto de género, el resultado lo que confirma es que por ser mujer se tiene un impacto diferenciado de ingresos en menos 41% con respecto a lo que perciben los hombres.

Y, con relación a los resultados comparativos del modelo en sus dos expresiones econométricas se tiene:

Tabla 38

Resultados comparativos de las regresiones de los modelos (2) y (3)

Modelo (2)					Modelo (3)				
Regresión lineal	Obs	=	56,086		Regresión lineal	Obs	=	56,086	
	F (4, 56081)	=	4860.6			F (5, 56080)	=	4428.64	
	Prob > F	=	0.0000			Prob > F	=	0.0000	
	R - squared	=	0.2699			R - squared	=	0.2862	
	Root MSE	=	.98468			Root MSE	=	.97366	

logIng	Coef	Robust Std. Err	t	P > t	logIng	Coef	Robust Std. Err	t	P > t
Educ	.1108613	.0010065	110.15	0.000	Educ	.0894523	.0011848	75.50	0.000
Edad	.0701489	.0014518	48.32	0.000	Edad	.0774528	.0014658	52.84	0.000
Edad2	-.0007134	.0000158	-45.03	0.000	Edad2	-.0007834	.0000159	-49.20	0.000
Sexo	-.397677	.0086205	-46.13	0.000	Sexo	-.4111998	.0085365	-48.17	0.000
_cons	4.118616	.0330355	124.67	0.000	TICs	.3959572	.0107496	36.83	0.000
					_cons	4.043689	.0332553	121.60	0.000

Nota: procesamiento propio en base a la encuesta de datos enaho 2019, resultado en anexo 5.

Se observa mejores resultados en el modelo (3) en el R cuadrado ajustado.

VI. CONCLUSIONES

En el presente trabajo de investigación se ha llegado a las siguientes conclusiones:

Primera: el nivel educativo alcanzado por las personas, en especial los estudiantes universitarios impactan positivamente en sus ingresos futuros, el impacto es de 11% en los ingresos.

Segunda: la experiencia de trabajo en el Perú, impacta positivamente en los ingresos que se obtiene por la actividad laboral, el impacto cuantificado es de 7% en los ingresos.

Tercera: el hecho de ser mujer o hombre impacta en los ingresos de forma diferenciada, ser mujer en el Perú es obtener menos 39.8% en los ingresos comparados a lo que obtienen los hombres.

Cuarta: en la segunda etapa del modelo conceptual se concluye que si aumenta en un año la educación, el impacto en los ingresos es el aumento en 9%, a más experiencia esto impacta en los ingresos en 8%, por ser mujer los ingresos son menores en 41% respecto a lo que perciben los hombres y si utilizan TICs esto impacta positivamente en sus ingresos, en 40%, que si no lo utilizan.

VII. RECOMENDACIONES

Se plantean las siguientes recomendaciones:

Profundizar las investigaciones en relación a los niveles educativos que la población universitaria logra, no se cuenta con información de detalle que se requiere para este tipo de investigaciones, la última encuesta data del 2015, pero es fundamental para la toma de decisiones en materia de docencia y gestión universitaria.

Se recomienda utilizar nuevas metodologías para el estudio del impacto de la experiencia de trabajo en los ingresos laborales, la universidad, la docencia y la investigación universitaria son quienes deben atender esta necesidad.

El problema de género en las remuneraciones y en el desarrollo de todas las actividades económicas en el Perú deberían de tener una mayor atención y para ello se recomienda a las universidades profundizar en la investigación académica sobre este tema en particular y a las instituciones públicas para efectos de mejorar las políticas públicas sobre el tema de género.

Se recomienda modificar las mallas curriculares en la formación profesional de los alumnos universitarios incluyendo el fortalecimiento en la formación como cursos transversales y obligatorios el uso de herramientas y software técnicos.

REFERENCIAS

- Alejo, J. (Setiembre de 2018). Educación y Desigualdad Salarial. *Investigaciones de Economía de la Educación, No. 13*, 293 - 314.
- Alejo, J., & Funes Leal, V. (2021). *Ecuaciones de Mincer de parejas bajo un esquema de selección muestral bivariada. Una aplicación al caso Argentino*. Uruguay: Instituto de Economía, Universidad de la República.
- Alshyab, N., Sandri, S., & Abu-Lila, Z. (2018). Differences in Returns to Education in Jordan: Gender and Public-private Sector Analysis. *55*(1).
- BCRP. (24 de Enero de 2022). <https://www.bcrp.gob.pe>. Obtenido de <https://www.bcrp.gob.pe/publicaciones/glosario>: <https://www.bcrp.gob.pe/publicaciones/glosario/s.html>
- Becker, G. S. (1975). *Human Capital, A Theoretical and Empirical Analysis, With Special Reference To Education*. New York: National Bureau of Economic Research.
- Becker, G. S. (1984). *El capital humano: Un análisis teórico y empírico referido fundamentalmente a la educación*. España: Alianza Editorial.
- Brassiolo, P., Nahirñak, P., & Ruffo, H. (2006). *Uso y adopción de tecnología informática en el mercado laboral de Argentina*. Buenos Aires, Argentina: Instituto de Estudios sobre la Realidad Argentina y Latinoamericana (IERAL).
- Castillo Robayo, C. D., Da Silva Bichara, J., & Pérez Trujillo, M. (2017). Retornos salariales para Colombia: un análisis cuantílico. *36*(63).
- Del Campo Villares, O., & Salcines Cristal, V. (2008). El valor económico de la educación a través del pensamiento económico en el Siglo XX. *Revista de la educación superior Vol. XXXVII, No. 147*, 45 - 61. Obtenido de http://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S0185-27602008000300004

- Félix Verduzco, G., & Torres García, A. J. (2018). Prima salarial al uso de computadora en el trabajo. Evidencia de microdatos para México. *Trimestre Económico*, 137 - 168.
- Galassi, G. L., & Andrada, M. J. (2009). *La Relación entre Educación e Ingresos: Ecuaciones de Mincer por Regiones Geográficas de Argentina*. San Fernando del Valle de Catamarca: Asociación de Estudios de Población de la Argentina.
- Gálvez Castex, P., & Valdés González, G. (2019). El retorno de la educación superior en Chile. *44*(8).
- Gil León, J. M., Casas Herrera, J. A., & Lemus Vergara, A. Y. (2020). ¿Es rentable la formación universitaria en Colombia?: una estimación. *10*(2).
- Gómez, M. C. (2018). Returns to Education and Skill Premiums: Estimation and Biases Associated with the Case of Argentina. (30).
- Hernández-Sampieri, R., & Mendoza Torres, C. P. (2018). *Metodología de la investigación: las rutas cuantitativas, cualitativa y mixta*. Ciudad de México: McGraw Hill Education.
- INEI. (2020). *Encuesta nacional de hogares 2019, calidad de la encuesta*. Lima, Perú: Instituto Nacional de estadística e informática.
- INEI. (2020). *Encuesta nacional de hogares 2019, diccionario*. Lima, Perú: Instituto Nacional de Estadística e Informática - INEI.
- INEI. (26 de Enero de 2022). <http://inei.inei.gob.pe>. Obtenido de <http://inei.inei.gob.pe/microdatos/>
- INEI. (5 de Enero de 2022). www.inei.gob.pe. Obtenido de www.inei.gob.pe/bases-de-datos/: <https://www.inei.gob.pe/bases-de-datos/>
- Instituto Nacional de Estadística e Informática. (2015). *Encuesta Nacional a Egresados Universitarios y Universidades, 2014*. Lima, Perú: Instituto Nacional de Estadística e Informática.
- Lucero, M. F. (2019). Rendimiento de la educación en Ecuador. (6).

- Mincer, J. (1981). *Human capital and economic growth*. Cambridge, EE.UU: National Bureau of economic research.
- Modrego Rico, A. (1989). La financiación de la inversión en capital humano. *Revista de economía (Madrid)*, 48 - 51.
- Montenegro, C. E., & Patrinos, H. A. (2014). *Comparable Estimates of Returns to Schooling Around the World*. Washington, DC: World Bank Group, Policy Research Working Paper, 7020.
- Mtpe, Minedu e Ipae. (2015). *Ponte en Carrera.pe*. Obtenido de Ponte en Carrera.pe: <https://www.ponteencarrera.pe/pec-portal-web/inicio/sobre-nosotros>
- OIT. (2019). *La brecha salarial entre hombres y mujeres en América Latina*. Oficina regional, Lima, Perú: OIT Américas.
- Ons Cappa, M., Sánchez Ollero, J., & García Pozo, A. (2020). Diferencias de género en los rendimientos del capital humano en el sector de la hostelería en España. (No.19).
- Pérez, J. E., Fuentes, H. J., & Contreras, L. E. (2020). Discriminación de género en el mercado laboral colombiano. *16(1)*.
- Ricardo Barreto, C., & Iriarte Diezgranados, F. (2017). *Las TIC en la educación superior, experiencias de innovación*. Barranquilla, Colombia: Universidad del Norte Editorial.
- Schultz, T. W. (1999). La inversión en capital humano. En M. F. Enguita (editor), *Sociología de la Educación*. Barcelona España: Ariel S.A.
- Sieber, S., Valor, J., & Porta, V. (2006). *Los sistemas de información en la empresa actual*. España: McGraw-Hill / Interamericana.
- Suárez Díaz, E. A., & Fuentes López, H. J. (2020). Análisis de los retornos de la educación para los jefes de hogar en Colombia. *15(2)*.
- Superintendencia Nacional de Educación Superior Universitaria. (2020). *II Informe bienal sobre la realidad universitaria en el Perú*. Lima, Perú: Sunedu.

- Torres García, A. J., & Ochoa Adame, G. L. (2018). Desigualdad salarial asociada al uso de las TIC en México: Un análisis por ocupaciones. *Cuadernos de Economía*, 37(74), 353 - 390.
- UCV. (23 de Mayo de 2017). Resolución de consejo universitario. *Código de ética en la investigación de la Universidad César Vallejo*. Trujillo, La Libertad, Perú: UCV.
- Unesco. (2016). *Tecnologías digitales al servicio de la calidad educativa*. Santiago de Chile: Unesco.
- Ventura Neyra, E. (2012). *Análisis de la heterogeneidad en los retornos a la educación en función del gasto público regional*. Lima - Perú: Pontificia Universidad Católica del Perú.
- Vicéns Otero, J., & Sánchez Reyes, B. (2012). Regresión Cuantílica: estimación y contrastes. (21).
- Villarreal Peralta, E. M. (2018). Endogeneidad de los rendimientos educativos en México. *Tomo 26(No.51)*.
- Wooldridge, J. M. (2010). *Introducción a la econometría, un enfoque moderno*. México D.F.: Cengage Learning.
- Yamada, G., & Castro, J. F. (2010). *Educación superior e ingresos laborales: Estimaciones paramétricas y no paramétricas de la rentabilidad por niveles y carreras en el Perú*. Lima, Perú: Centro de Investigación de la Universidad del Pacífico. Obtenido de <https://repositorio.up.edu.pe/bitstream/handle/11354/359/DD1006.pdf?sequence=1&isAllowed=y>
- Yamada, G., Lavado, P., & Oviedo, N. (2016). *La evidencia de rendimientos de la educación superior a partir de "Ponte en Carrera"*. Lima, Perú: Universidad del Pacífico.

Anexo 1

Matriz de Consistencia

PROBLEMAS	OBJETIVOS	HIPOTESIS	VARIABLES
Primera etapa			Variables independientes
<p>Problema general ¿Cuál es el impacto del nivel educativo logrado, la experiencia que tienen, y el género, de los egresados universitarios, en los ingresos que perciben en su actividad laboral?</p>	<p>Objetivo general Determinar el impacto, del nivel educativo logrado, la experiencia que tienen, y el género, de los egresados universitarios, en el nivel de ingresos que perciben en su actividad laboral.</p>	<p>Hipótesis general Existe un impacto positivo, del nivel educativo logrado y la experiencia de los egresados universitarios, en el nivel de ingresos de estos en su actividad laboral, y un impacto diferenciado entre el género de los egresados universitarios con el nivel de ingresos de estos en su actividad laboral.</p>	<ol style="list-style-type: none"> 1. Nivel educativo 2. Experiencia 3. Género
Segunda etapa			Variable dependiente
<p>Problema general ¿Cuál es el impacto del nivel educativo logrado, el manejo de las TICs, la experiencia que tienen, y el género, de los egresados universitarios, en los ingresos que perciben en su actividad laboral?</p>	<p>Objetivo general Determinar el impacto, del nivel educativo logrado, el manejo de las TICs, la experiencia que tienen, y el género, de los egresados universitarios, en el nivel de ingresos que perciben en su actividad laboral.</p>	<p>Hipótesis general Existe un impacto positivo, del nivel educativo logrado, el manejo de las TICs y la experiencia, de los egresados universitarios, en el nivel de ingresos de estos en su actividad laboral, y un impacto diferenciado entre el género de los egresados universitarios con el nivel de ingresos de estos en su actividad laboral.</p>	<ol style="list-style-type: none"> 1. Nivel de ingresos

Anexo 2

Matriz de operacionalización de las variables

VARIABLES DE ESTUDIO	DEFINICIÓN CONCEPTUAL	DEFINICIÓN OPERACIONAL	DIMENSIONES	INDICADORES	ESCALA DE MEDICIÓN
Variable dependiente: INGRESOS	La variable "ingresos denota a los salarios, tomados como ingresos por hora o como ingresos mensuales" expresados en logaritmo (Suarez Díaz y Fuentes López, 2020)	Son los ingresos totales percibidos en un determinado periodo, más los pagos percibidos en especie, más el ingreso por labor independiente si lo percibe y más el valor de los productos utilizados para su consumo, expresado mensualmente	Logaritmo de los ingresos	Log (Ing)	Razón
Variable independiente: NIVEL EDUCATIVO	La variable nivel educativo se basa en el concepto del "proceso de formación del individuo que genera un aumento en la capacidad productiva (Modrego, 1992, 48-51)" citado en (Del Campo Villares y Salcines Cristal, 2008)	Es el número de años invertidos en el proceso de su formación educativa que comprende la formación básica regular y la formación superior	Años de educación	Número de años de educación	Razón
Variable independiente: EXPERIENCIA	La variable experiencia se enfoca bajo el concepto de que es "la medida que aumenta el nivel de entrenamiento y especialización que hay en una empresa que aumentarán también las diferencias salariales entre el grueso de la población laboral" (Mincer, 1981)	Para esta variable, que indica la experiencia laboral del individuo, se consideró los años del trabajador como una variable proxy de la experiencia laboral. Entre otras investigaciones que asumen esta misma metodología por ejemplo (Gálvez Castex y Valdés González, 2019)	Años cumplidos	Número de años cumplidos	Razón
Variable independiente: GÉNERO	La variable género si bien es cierto que está referida al hecho de ser mujer o hombre, se refiere al concepto propuesto por la OIT "trabajo decente para todas las personas" (OIT, 2019)	Variable dicotómica que se refiere al hecho de ser mujer o de ser hombre, por tanto, a efectos de su operacionalización consideramos como sexo.	Sexo	0 = hombre 1 = mujer	Nominal
Variable independiente: TICs	La variable TICs se refiere a "las herramientas para la gestión de la información que permiten y proporcionan soluciones y apoyo a la toma de decisiones" (Sieber, Valor y Porta, 2006)	Para esta variable, se considera si conoce y utiliza o no, la computadora, en sus labores	TICs variable binaria	0 = no utiliza 1 = si utiliza	Nominal

Anexo 3
Listado de temas de estudio del Enaho 2019

Nro	Año	Período	Código Encuesta	Encuesta	Código Módulo	Módulo
1	2019	55	687	Condiciones de Vida y Pobreza - ENAHO	1	Características de la Vivienda y del Hogar
2	2019	55	687	Condiciones de Vida y Pobreza - ENAHO	2	Características de los Miembros del Hogar
3	2019	55	687	Condiciones de Vida y Pobreza - ENAHO	3	Educación
4	2019	55	687	Condiciones de Vida y Pobreza - ENAHO	4	Salud
5	2019	55	687	Condiciones de Vida y Pobreza - ENAHO	5	Empleo e Ingresos
6	2019	55	687	Condiciones de Vida y Pobreza - ENAHO	7	Gastos en Alimentos y Bebidas (Módulo 601)
7	2019	55	687	Condiciones de Vida y Pobreza - ENAHO	8	Instituciones Benéficas
8	2019	55	687	Condiciones de Vida y Pobreza - ENAHO	9	Mantenimiento de la Vivienda
9	2019	55	687	Condiciones de Vida y Pobreza - ENAHO	10	Transportes y Comunicaciones
10	2019	55	687	Condiciones de Vida y Pobreza - ENAHO	11	Servicios a la Vivienda
11	2019	55	687	Condiciones de Vida y Pobreza - ENAHO	12	Esparcimiento , Diversion y Servicios de Cultura
12	2019	55	687	Condiciones de Vida y Pobreza - ENAHO	13	Vestido y Calzado
13	2019	55	687	Condiciones de Vida y Pobreza - ENAHO	15	Gastos de Transferencias
14	2019	55	687	Condiciones de Vida y Pobreza - ENAHO	16	Muebles y Enseres
15	2019	55	687	Condiciones de Vida y Pobreza - ENAHO	17	Otros Bienes y Servicios
16	2019	55	687	Condiciones de Vida y Pobreza - ENAHO	18	Equipamiento del Hogar
17	2019	55	687	Condiciones de Vida y Pobreza - ENAHO	22	Producción Agrícola
18	2019	55	687	Condiciones de Vida y Pobreza - ENAHO	23	Subproductos Agrícolas
19	2019	55	687	Condiciones de Vida y Pobreza - ENAHO	24	Producción Forestal
20	2019	55	687	Condiciones de Vida y Pobreza - ENAHO	25	Gastos en Actividades Agrícolas y/o Forestales
21	2019	55	687	Condiciones de Vida y Pobreza - ENAHO	26	Producción Pecuaria
22	2019	55	687	Condiciones de Vida y Pobreza - ENAHO	27	Subproductos Pecuarios
23	2019	55	687	Condiciones de Vida y Pobreza - ENAHO	28	Gastos en Actividades Pecuarias
24	2019	55	687	Condiciones de Vida y Pobreza - ENAHO	34	Sumarias (Variables Calculadas)
25	2019	55	687	Condiciones de Vida y Pobreza - ENAHO	37	Programas Sociales (Miembros del Hogar)
26	2019	55	687	Condiciones de Vida y Pobreza - ENAHO	77	Ingresos del Trabajador Independiente
27	2019	55	687	Condiciones de Vida y Pobreza - ENAHO	78	Bienes y Servicios de Cuidados Personales
28	2019	55	687	Condiciones de Vida y Pobreza - ENAHO	84	Participación Ciudadana
29	2019	55	687	Condiciones de Vida y Pobreza - ENAHO	85	Gobernabilidad, Democracia y Transparencia

Anexo 4

Programación en el programa Stata

Contains data from C:\Users\Angel\OneDrive\Escritorio\Prueba_para_tesis_resultados\BaseD
> a123.dta

obs: 128,276
vars: 33 20 Feb 2022 16:10

variable name	storage type	display format	value label	variable label
conglome	str6	%6s		número de conglomerado
vivienda	str3	%3s		número de selección de la vivienda
hogar	str2	%2s		número secuencial del hogar
codperso	str2	%2s		número de orden de la persona
ubigeo	str6	%6s		ubicación geográfica
dominio	byte	%8.0g	dominio	dominio geográfico
estrato	byte	%47.0g	estrato	estrato geográfico
p207	byte	%8.0g	p207	sexo
Edad	byte	%8.0g		¿qué edad tiene en años cumplidos ? (en años)
facpob07	float	%9.0g		factor de expansión anual de población proyecciones cpv-2007
Sexo	float	%9.0g	Sexo_label	Sexo del entrevistado
nconglome	str6	%6s		número de conglomerado (proveniente del marco)
codinfor	str2	%2s		código de informante del capítulo 300
p301a	byte	%36.0g	p301a	¿cuál es el último año o grado de estudios y nivel que aprobó? - nivel
p301b	byte	%8.0g		¿cuál es el último año o grado de estudios y nivel que aprobó? - año
p301c	byte	%8.0g		¿cuál es el último año o grado de estudios y nivel que aprobó? - grado
p301d	byte	%8.0g	p301d	¿cuál es el último año o grado de estudios y nivel que aprobó? - centro de estudio
imputado	byte	%24.0g	imputado	conglomerado - (imputación hot-deck)
factor07	float	%9.0g		factor de expansión anual proyecciones cpv-2007
factora07	float	%9.0g		factor de población ajustado por grupos de edad
Educ	float	%9.0g		Grado de estudios y nivel que aprobo
ocu500	byte	%8.0g	ocu500	indicador de la pea
ticuest01a	byte	%8.0g	ticuest01a	origen de cuestionario
d529t	long	%12.0g		(deflactado, anualizado) en su ocupación principal, ¿con qué frecuencia y en cuánta
d536	int	%8.0g		(deflactado, anualizado) ¿en cuánto estima ud. el valor de los productos utiliza
i530a	double	%10.0g		(imputado, deflactado, anualizado) en la ocupación principal ¿cuál fue la ganancia
i524a1	double	%10.0g		(imputado, deflactado, anualizado) ¿cuánto fue su ingreso total en el(la) ... an
fac500a	float	%9.0g		factor de expansión de empleo/ingresos proyecciones cpv-2007
inganual	float	%9.0g		Ingreso mensual
Ing	float	%9.0g		Ingreso mensual
_merge	byte	%23.0g	_merge	Logaritmo del Ingreso mensual
logIng	float	%9.0g		Logaritmo del Ingreso mensual
Edad2	float	%9.0g		Edad al cuadrado

Sorted by: conglome vivienda hogar codperso ubigeo dominio estrato

Anexo 5
Resultados de las corridas en el Stata

Ing Ingreso mensual

type: numeric (float)

range: [0,26018.334] units: 1.000e-07
unique values: 26,461 missing .: 63,322/128,276

mean: 999.245
std. dev: 1270.1

percentiles: 10% 25% 50% 75% 90%
 0 177.333 656.583 1332.75 2309.42

Educ Grado de estudios y nivel que aprobo

type: numeric (float)

range: [0,18] units: 1
unique values: 19 missing .: 11,876/128,276

mean: 7.65103
std. dev: 5.00579

percentiles: 10% 25% 50% 75% 90%
 0 3 8 11 14

Edad ¿qué edad tiene en años cumplidos ? (en años)

type: numeric (byte)

range: [0,98] units: 1
unique values: 99 missing .: 3,297/128,276

mean: 33.7339
std. dev: 22.8636

percentiles: 10% 25% 50% 75% 90%
 6 14 31 51 67

Grado de estudios y nivel que aprobo	Sexo del entrevistado		Total
	Hombre	Mujer	
0	5,723	9,275	14,998
1	1,938	2,466	4,404
2	2,552	3,124	5,676
3	2,815	3,134	5,949
4	2,195	2,345	4,540
5	1,656	1,697	3,353
6	6,552	6,571	13,123
7	2,159	1,963	4,122
8	2,606	2,544	5,150
9	2,628	2,457	5,085
10	1,802	1,791	3,593
11	11,841	9,740	21,581
12	1,787	1,810	3,597
13	2,243	2,110	4,353
14	3,610	3,825	7,435
15	747	730	1,477
16	3,516	3,774	7,290
17	280	316	596
18	46	32	78
Total	56,696	59,704	116,400

Variable	Obs	Mean	Std. Dev.	Min	Max
logIng	56,101	6.520618	1.152358	.5596158	10.16656
Educ	116,400	7.651031	5.005789	0	18
Edad	124,979	33.73394	22.86356	0	98
Edad2	124,979	1660.717	1845.779	0	9604
Sexo	124,979	.511718	.4998647	0	1

	logIng	Educ	Edad	Edad2
logIng	1.0000			
Educ	0.4541*	1.0000		
Edad	-0.1147*	0.1086*	1.0000	
Edad2	-0.1596*	-0.0383*	0.9590*	1.0000

Source	SS	df	MS	Number of obs	=	56,086
Model	20104.6942	4	5026.17355	F(4, 56081)	=	5183.77
Residual	54376.0595	56,081	.969598607	Prob > F	=	0.0000
Total	74480.7537	56,085	1.32799775	R-squared	=	0.2699
				Adj R-squared	=	0.2699
				Root MSE	=	.98468

logIng	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
Educ	.1108613	.0009929	111.65	0.000	.1089152	.1128075
Edad	.0701489	.0013296	52.76	0.000	.0675429	.072755
Edad2	-.0007134	.0000142	-50.19	0.000	-.0007413	-.0006856
Sexo	-.397677	.0084339	-47.15	0.000	-.4142074	-.3811465
_cons	4.118616	.0309389	133.12	0.000	4.057975	4.179256

Breusch-Pagan / Cook-Weisberg test for heteroskedasticity

Ho: Constant variance

Variables: fitted values of logIng

chi2(1) = 914.52

Prob > chi2 = 0.0000

```
. vif
```

Variable	VIF	1/VIF
Edad2	26.99	0.037052
Edad	26.46	0.037797
Educ	1.21	0.826676
Sexo	1.00	0.996356
Mean VIF	13.91	

```
. tab TICs Sexo
```

Uso de TICs	Sexo del entrevistado		Total
	Hombre	Mujer	
No	29,719	34,363	64,082
Si	14,635	13,344	27,979
Total	44,354	47,707	92,061

```
Breusch-Pagan / Cook-Weisberg test for heteroskedasticity
Ho: Constant variance
Variables: fitted values of logIng

chi2(1)      = 1218.58
Prob > chi2  = 0.0000
```

Variable	VIF	1/VIF
Edad2	27.51	0.036346
Edad	27.10	0.036905
Educ	1.66	0.602367
TICs	1.50	0.667870
Sexo	1.01	0.994308
Mean VIF	11.75	

Linear regression		Number of obs	=	56,086	
		F(4, 56081)	=	4860.60	
		Prob > F	=	0.0000	
		R-squared	=	0.2699	
		Root MSE	=	.98468	
logIng	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]
Educ	.1108613	.0010065	110.15	0.000	.1088887 .112834
Edad	.0701489	.0014518	48.32	0.000	.0673034 .0729945
Edad2	-.0007134	.0000158	-45.03	0.000	-.0007445 -.0006824
Sexo	-.397677	.0086205	-46.13	0.000	-.4145733 -.3807806
_cons	4.118616	.0330355	124.67	0.000	4.053866 4.183366

Grado de estudios y nivel que aprobo	Uso de TICs		Total
	No	Si	
0	6,433	9	6,442
1	2,106	5	2,111
2	3,129	21	3,150
3	3,403	28	3,431
4	2,134	32	2,166
5	1,063	37	1,100
6	10,758	314	11,072
7	2,115	505	2,620
8	3,342	1,443	4,785
9	3,502	1,572	5,074
10	1,936	1,657	3,593
11	16,315	5,266	21,581
12	1,297	2,300	3,597
13	1,191	3,162	4,353
14	3,143	4,292	7,435
15	301	1,176	1,477
16	1,732	5,558	7,290
17	75	521	596
18	6	72	78
Total	63,981	27,970	91,951

Variable	Obs	Mean	Std. Dev.	Min	Max
logIng	56,101	6.520618	1.152358	.5596158	10.16656
Educ	116,400	7.651031	5.005789	0	18
Edad	124,979	33.73394	22.86356	0	98
Edad2	124,979	1660.717	1845.779	0	9604
Sexo	124,979	.511718	.4998647	0	1

	logIng	Educ	Edad	Edad2	Sexo	TICs
logIng	1.0000					
Educ	0.4541*	1.0000				
Edad	-0.1147*	0.1086*	1.0000			
Edad2	-0.1596*	-0.0383*	0.9590*	1.0000		
Sexo	-0.1814*	-0.0692*	0.0294*	0.0262*	1.0000	
TICs	0.3271*	0.5096*	-0.3773*	-0.3474*	-0.0546*	1.0000

Source	SS	df	MS	Number of obs	=	56,086
Model	21315.2175	5	4263.04349	F(5, 56080)	=	4496.74
Residual	53165.5362	56,080	.948030246	Prob > F	=	0.0000
				R-squared	=	0.2862
				Adj R-squared	=	0.2861
Total	74480.7537	56,085	1.32799775	Root MSE	=	.97367

logIng	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
Educ	.0894523	.0011502	77.77	0.000	.087198	.0917067
Edad	.0774528	.0013305	58.21	0.000	.0748449	.0800606
Edad2	-.0007834	.0000142	-55.21	0.000	-.0008112	-.0007556
Sexo	-.4111998	.0083481	-49.26	0.000	-.4275621	-.3948374
TICs	.3959572	.0110808	35.73	0.000	.3742387	.4176758
_cons	4.043689	.0306647	131.87	0.000	3.983586	4.103792

stats	Ing	Educ	Edad
N	64954	116400	124979
max	26018.33	18	98
min	0	0	0
range	26018.33	18	98
mean	999.2453	7.651031	33.73394
sd	1270.104	5.005789	22.86356
cv	1.271063	.6542633	.6777613

Linear regression

Number of obs = 56,086
 F(5, 56080) = 4428.64
 Prob > F = 0.0000
 R-squared = 0.2862
 Root MSE = .97367

logIng	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
Educ	.0894523	.0011848	75.50	0.000	.0871302	.0917745
Edad	.0774528	.0014658	52.84	0.000	.0745799	.0803257
Edad2	-.0007834	.0000159	-49.20	0.000	-.0008146	-.0007522
Sexo	-.4111998	.0085365	-48.17	0.000	-.4279314	-.3944681
TICs	.3959572	.0107496	36.83	0.000	.374888	.4170265
_cons	4.043689	.0332553	121.60	0.000	3.978509	4.10887

Grado de estudios y nivel que aprobo	Sexo del entrevistado		Total
	Hombre	Mujer	
0	238,703.5	780,592	1019295.5
1	128,991.6	253,880.6	382,872.1
2	229,639.4	348,031.5	577,670.9
3	305,993.6	388,039.9	694,033.4
4	194,839.7	238,411.1	433,250.8
5	118,700.5	137,376.6	256,077.1
6	1179824.1	1249111.9	2,428,936
7	347,865.3	304,166.86	652,032.16
8	667,723.4	641,780.3	1309503.7
9	745,145.5	679,573.91	1424719.4
10	534,019.7	529,786.8	1063806.4
11	3804250.3	3015702.9	6819953.2
12	582,745.5	584,679.94	1167425.4
13	707,826.6	667,359.61	1375186.2
14	1215290.1	1233566.9	2,448,857
15	238,464.7	233,318.9	471,783.6
16	1053819.1	1071562.8	2125381.9
17	78,891.36	80,687.13	159,578.5
18	15,338.26	10,507.51	25,845.77
Total	12,388,072	12,448,137	24,836,209

logIng		Robust		t	P> t	[95% Conf. Interval]	
	Coef.	Std. Err.					
Educ	.1108613	.0010065	110.15	0.000	.1088887	.112834	
Edad	.0701489	.0014518	48.32	0.000	.0673034	.0729945	
Edad2	-.0007134	.0000158	-45.03	0.000	-.0007445	-.0006824	
Sexo	-.397677	.0086205	-46.13	0.000	-.4145733	-.3807806	
_cons	4.118616	.0330355	124.67	0.000	4.053866	4.183366	

Linear regression

Number of obs = 56,086
F(4, 56081) = 4860.60
Prob > F = 0.0000
R-squared = 0.2699
Root MSE = .98468

Linear regression

Number of obs = 56,086
F(5, 56080) = 4428.64
Prob > F = 0.0000
R-squared = 0.2862
Root MSE = .97367

logIng

Robust
Coef. Std. Err. t P>|t| [95% Conf. Interval]

Educ	.0894523	.0011848	75.50	0.000	.0871302	.0917745
Edad	.0774528	.0014658	52.84	0.000	.0745799	.0803257
Edad2	-.0007834	.0000159	-49.20	0.000	-.0008146	-.0007522
Sexo	-.4111998	.0085365	-48.17	0.000	-.4279314	-.3944681
TICs	.3959572	.0107496	36.83	0.000	.374888	.4170265
_cons	4.043689	.0332553	121.60	0.000	3.978509	4.10887