

Universidad de Lima

Facultad de Ciencias Empresariales y Económicas

Carrera de Economía



**IMPACTO DE LA SALUD EN LA
PRODUCTIVIDAD EN EL PERÚ, 2019:
ANÁLISIS DESAGREGADO POR DOMINIOS
GEOGRÁFICOS, RANGOS DE EDAD,
NIVELES EDUCATIVOS Y QUINTILES DE
INGRESO**

Tesis para optar el Título Profesional de Economista

Aldo Franco Obregón Donayre

Código 20161010

Schilder Lenin Rondan Bravo

Código 20161287

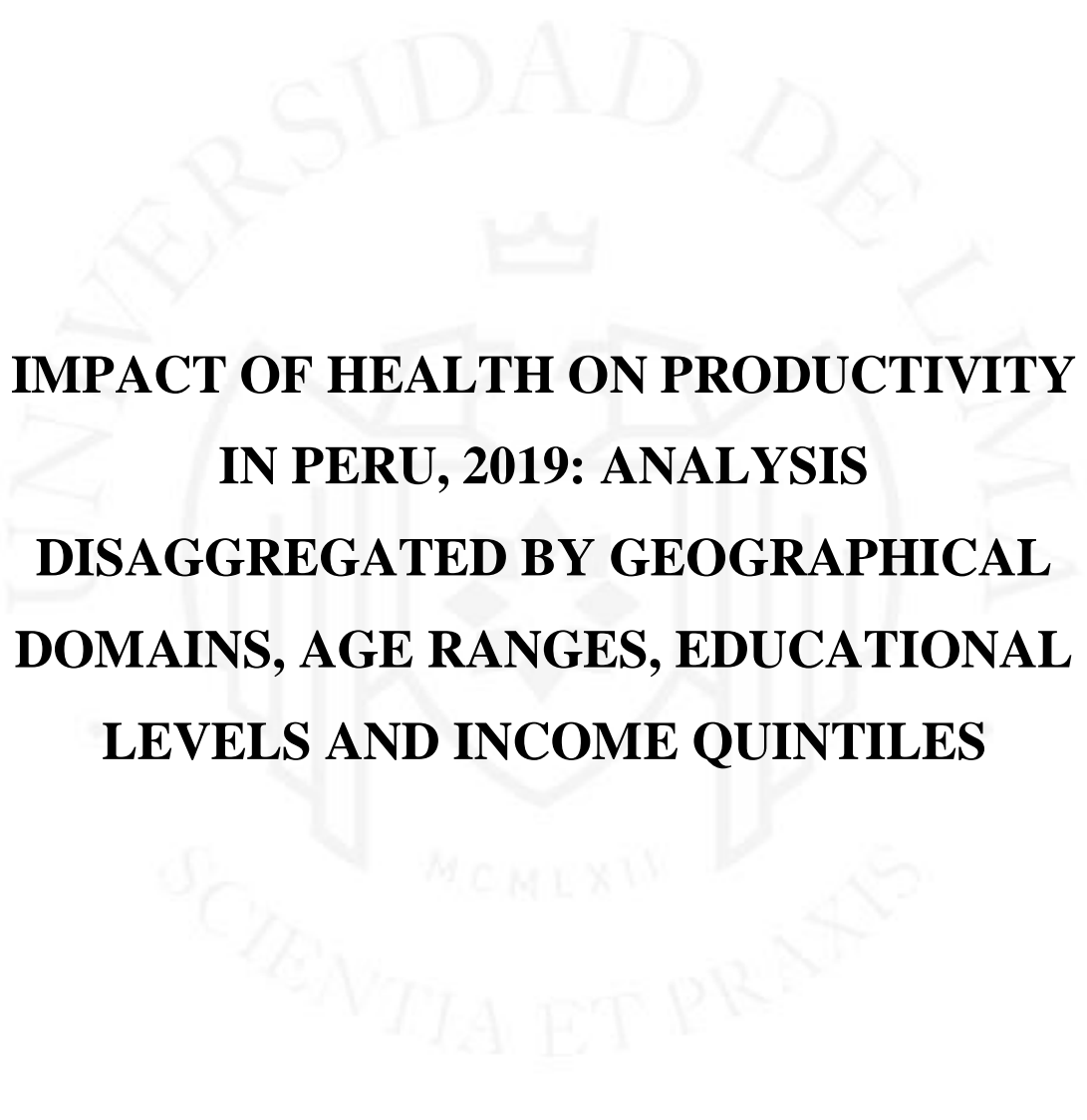
Asesora

Rosa Luz Durán Fernández

Lima – Perú

Febrero de 2023





**IMPACT OF HEALTH ON PRODUCTIVITY
IN PERU, 2019: ANALYSIS
DISAGGREGATED BY GEOGRAPHICAL
DOMAINS, AGE RANGES, EDUCATIONAL
LEVELS AND INCOME QUINTILES**

TABLA DE CONTENIDO

RESUMEN	XI
ABSTRACT	XIII
INTRODUCCIÓN.....	1
CAPÍTULO I: ANTECEDENTES	6
CAPÍTULO II: REVISIÓN DE LITERATURA	13
2.1 Marco teórico	13
2.2 Literatura empírica.....	16
CAPÍTULO III: METODOLOGÍA	24
3.1 Datos y variables	24
3.2 Operacionalización de las variables.....	26
3.3 Propuesta de estimación empírica.....	28
3.4 Estadísticas Descriptivas	31
CAPÍTULO IV: ANÁLISIS DE RESULTADOS	54
CAPÍTULO V: CONCLUSIONES	69
CAPÍTULO VI: RECOMENDACIONES	71
REFERENCIAS	73
FUENTES DE DATOS.....	76

ÍNDICE DE TABLAS

Tabla 3.1: Operacionalización de las variables	27
Tabla 3.2: Características de la muestra de encuestados desagregado por el indicador de salud y género	32
Tabla 3.3: Características del hogar sobre la muestra de encuestados desagregado por el indicador de salud.....	35
Tabla 3.4: Ingreso promedio desagregado por dominios geográficos, indicador de salud y sexo	37
Tabla 3.5: Número de individuos enfermos y no enfermos desagregado por dominios geográficos.....	39
Tabla 3.6: Ingreso promedio desagregado por niveles educativos, indicador de salud y sexo	42
Tabla 3.7: Número de individuos enfermos y no enfermos desagregado por niveles educativos	44
Tabla 3.8: Ingreso promedio desagregado por quintiles de ingreso, indicador de salud y sexo	46
Tabla 3.9: Número de individuos enfermos y no enfermos desagregado por quintiles de ingreso.....	48
Tabla 3.10: Ingreso promedio desagregado por rangos de edad, indicador de salud y sexo	50
Tabla 3.11: Número de individuos enfermos y no enfermos desagregado por rangos de edad	52
Tabla 4.1: Resultado econométrico a nivel general mediante el modelo MCO corregido por la estimación en dos etapas de Heckman	54
Tabla 4.2: Resultado econométrico sobre los niveles educativos mediante el modelo MCO corregido por la estimación en dos etapas de Heckman.....	56

Tabla 4.3: Resultado econométrico sobre los dominios geográficos mediante el modelo MCO corregido por la estimación en dos etapas de Heckman.....	59
Tabla 4.4: Resultado econométrico sobre los quintiles de ingreso mediante el modelo MCO corregido por la estimación en dos etapas de Heckman.....	62
Tabla 4.5: Resultado econométrico sobre los quintiles de ingreso mediante el modelo MCO corregido por la estimación en dos etapas de Heckman.....	65



ÍNDICE DE FIGURAS

Figura 1.1: Evolución del número de hospitalizados en Perú desde el año 2007 al 2017	7
Figura 1.2: Evolución de la remuneración mínima vital en Perú desde 1990 hasta el 2017	9
Figura 1.3: Evolución del ingreso promedio mensual proveniente del trabajo por regiones del Perú desde 2007 hasta el 2018	10
Figura 1.4: Evolución de la población afiliada a seguro de salud según zona en Perú desde 2007 hasta el 2018.....	11
Figura 1.5: Número de habitantes por cada médico en Perú desde 2010 hasta el 2018. 12	
Figura 3.1: Ingreso promedio desagregado por dominios geográficos, indicador de salud y sexo	38
Figura 3.2: Número de individuos enfermos y no enfermos desagregado por dominios geográficos.....	41
Figura 3.3: Ingreso promedio desagregado por niveles educativos, indicador de salud y sexo	43
Figura 3.4: Número de individuos enfermos y no enfermos desagregado por niveles educativo.....	45
Figura 3.5: Ingreso promedio desagregado por quintiles de ingreso, indicador de salud y sexo	47
Figura 3.6: Número de individuos enfermos y no enfermos desagregado por quintiles de ingreso.....	49
Figura 3.7: Ingreso promedio desagregado por rangos de edad, indicador de salud y sexo	51
Figura 3.8: Número de individuos enfermos y no enfermos desagregado por rangos de edad	53

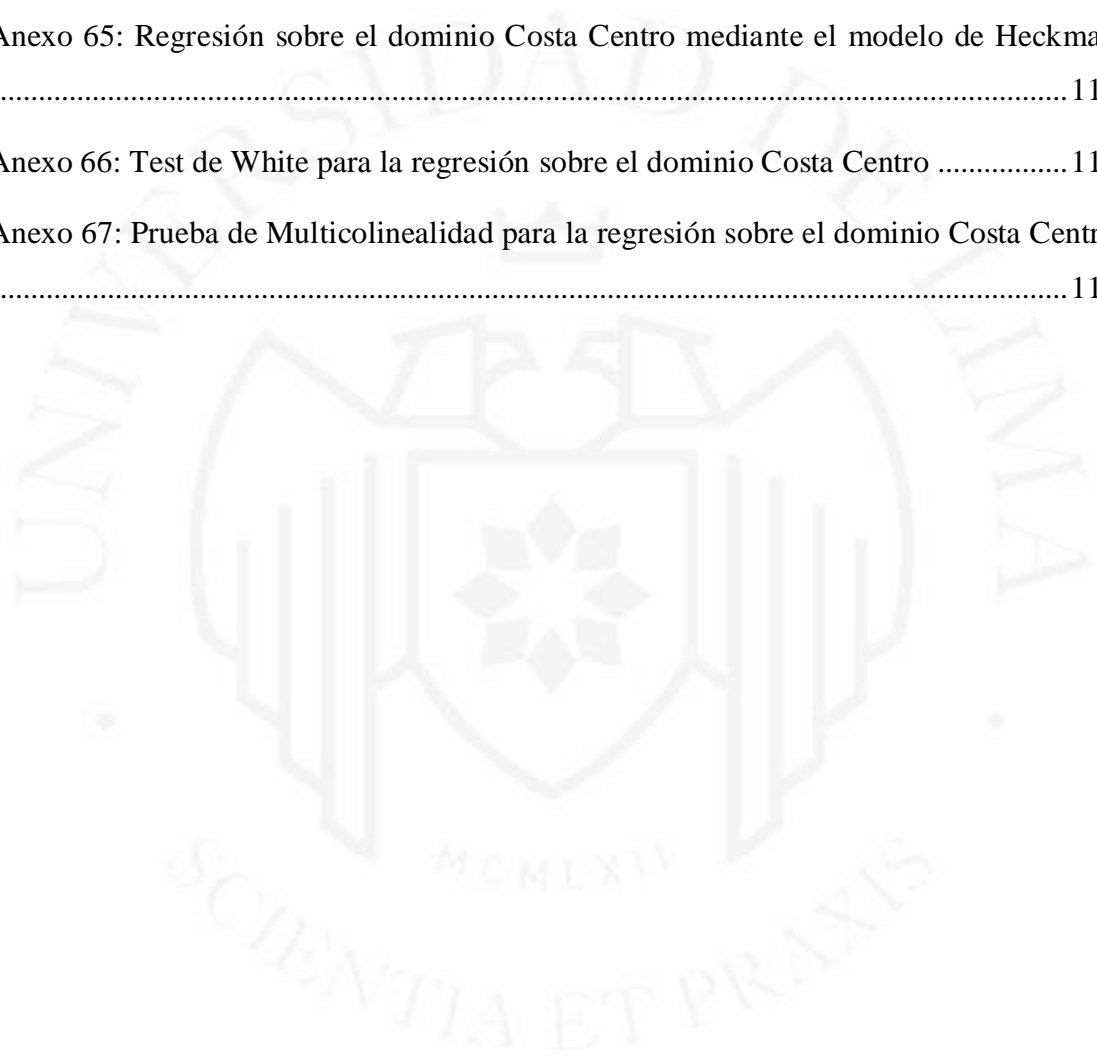
ÍNDICE DE ANEXOS

Anexo 1: Matriz de Consistencia	78
Anexo 2: Regresión General mediante el modelo de Heckman.....	81
Anexo 3: Test de White para la regresión general	81
Anexo 4: Prueba de Multicolinealidad para la regresión general	82
Anexo 5: Regresión sobre el rango de edad 18-24 mediante el modelo de Heckman....	82
Anexo 6: Test de White para la regresión sobre el rango de edad 18-24	83
Anexo 7: Prueba de Multicolinealidad para la regresión sobre el rango de edad 18-24	83
Anexo 8: Regresión sobre el rango de edad 25-34 mediante el modelo de Heckman....	84
Anexo 9: Test de White para la regresión sobre el rango de edad 25-34	84
Anexo 10: Prueba de Multicolinealidad para la regresión sobre el rango de edad 25-34	85
Anexo 11: Regresión sobre el rango de edad 35-44 mediante el modelo de Heckman .	85
Anexo 12: Test de White para la regresión sobre el rango de edad 35-44	86
Anexo 13: Prueba de Multicolinealidad para la regresión sobre el rango de edad 35-44	86
Anexo 14: Regresión sobre el rango de edad 45-59 mediante el modelo de Heckman .	87
Anexo 15: Test de White para la regresión sobre el rango de edad 45-59	87
Anexo 16: Prueba de Multicolinealidad para la regresión sobre el rango de edad 45-59	88
Anexo 17: Regresión sobre el rango de edad 60-70 mediante el modelo de Heckman .	88
Anexo 18: Test de White para la regresión sobre el rango de edad 60-70	89
Anexo 19: Prueba de Multicolinealidad para la regresión sobre el rango de edad 60-70	89

Anexo 20: Regresión sobre el primer quintil mediante el modelo de Heckman.....	90
Anexo 21: Test de White para la regresión sobre el primer quintil	90
Anexo 22: Prueba de Multicolinealidad para la regresión sobre el primer quintil	91
Anexo 23: Regresión sobre el segundo quintil mediante el modelo de Heckman.....	91
Anexo 24: Test de White para la regresión sobre el segundo quintil	92
Anexo 25: Prueba de Multicolinealidad para la regresión sobre el segundo quintil.....	92
Anexo 26: Regresión sobre el tercer quintil mediante el modelo de Heckman	93
Anexo 27: Test de White para la regresión sobre el tercer quintil.....	93
Anexo 28: Prueba de Multicolinealidad para la regresión sobre el tercer quintil	94
Anexo 29: Regresión sobre el cuarto quintil mediante el modelo de Heckman	94
Anexo 30: Test de White para la regresión sobre el cuarto quintil	95
Anexo 31: Prueba de Multicolinealidad para la regresión sobre el cuarto quintil	95
Anexo 32: Regresión sobre el quinto quintil mediante el modelo de Heckman	96
Anexo 33: Test de White para la regresión sobre el quinto quintil.....	96
Anexo 34: Prueba de Multicolinealidad para la regresión sobre el quinto quintil	97
Anexo 35: Regresión sobre el nivel educativo primaria mediante el modelo de Heckman	97
Anexo 36: Test de White para la regresión sobre el nivel educativo primaria.....	98
Anexo 37: Prueba de Multicolinealidad para la regresión sobre el nivel educativo primaria	98
Anexo 38: Regresión sobre el nivel educativo secundaria mediante el modelo de Heckman	99
Anexo 39: Test de White para la regresión sobre el nivel educativo secundaria.....	99
Anexo 40: Prueba de Multicolinealidad para la regresión sobre el nivel educativo secundaria	100
Anexo 41: Regresión sobre el nivel educativo superior mediante el modelo de Heckman	100

Anexo 42: Test de White para la regresión sobre el nivel educativo superior	101
Anexo 43: Prueba de Multicolinealidad para la regresión sobre el nivel educativo superior	101
Anexo 44: Regresión sobre el dominio Sierra Sur mediante el modelo de Heckman ..	102
Anexo 45: Test de White para la regresión sobre el dominio Sierra Sur.....	102
Anexo 46: Prueba de Multicolinealidad para la regresión sobre el dominio Sierra Sur	103
Anexo 47: Regresión sobre el dominio Sierra Norte mediante el modelo de Heckman	103
Anexo 48: Test de White para la regresión sobre el dominio Sierra Norte	104
Anexo 49: Prueba de Multicolinealidad para la regresión sobre el dominio Sierra Norte	104
Anexo 50: Regresión sobre el dominio Sierra Centro mediante el modelo de Heckman	105
Anexo 51: Test de White para la regresión sobre el dominio Sierra Centro	105
Anexo 52: Prueba de Multicolinealidad para la regresión sobre el dominio Sierra Centro	106
Anexo 53: Regresión sobre el dominio Selva mediante el modelo de Heckman.....	106
Anexo 54: Test de White para la regresión sobre el dominio Selva	107
Anexo 55: Prueba de Multicolinealidad para la regresión sobre el dominio Selva	107
Anexo 56: Regresión sobre el dominio Lima Metropolitana mediante el modelo de Heckman	108
Anexo 57: Test de White para la regresión sobre el dominio Lima Metropolitana	108
Anexo 58: Prueba de Multicolinealidad para la regresión sobre el dominio Lima Metropolitana	109
Anexo 59: Regresión sobre el dominio Costa Sur mediante el modelo de Heckman ..	109
Anexo 60: Test de White para la regresión sobre el dominio Costa Sur	110

Anexo 61: Prueba de Multicolinealidad para la regresión sobre el dominio Costa Sur	110
Anexo 62: Regresión sobre el dominio Costa Norte mediante el modelo de Heckman	111
Anexo 63: Test de White para la regresión sobre el dominio Costa Norte	111
Anexo 64: Prueba de Multicolinealidad para la regresión sobre el dominio Costa Norte	112
Anexo 65: Regresión sobre el dominio Costa Centro mediante el modelo de Heckman	112
Anexo 66: Test de White para la regresión sobre el dominio Costa Centro	113
Anexo 67: Prueba de Multicolinealidad para la regresión sobre el dominio Costa Centro	113



RESUMEN

El presente trabajo de investigación tiene como principal objetivo identificar el impacto de la salud sobre la productividad en el Perú para el año 2019, con respecto a 4 desagregaciones, estas son: niveles educativos, dominios geográficos, quintiles de ingreso y rangos de edad. Para tal fin se empleó un modelo econométrico de Mínimos Cuadrados Ordinarios (MCO) corregido por la estimación en dos etapas de Heckman, el cual tiene la ventaja de corregir el sesgo de selección que se presenta en la estimación de la ecuación de los salarios. El estudio abarca a 22,061 encuestados seleccionados de la Encuesta Nacional de Hogares (ENAHOG); para el análisis se agrupan las variables explicativas de acuerdo con las características individuales, variables del capital humano, variables sobre la vivienda y variables del mercado local. Asimismo, dado que la productividad no se puede medir de forma tradicional, el presente trabajo de investigación establece como indicador de productividad a la variable de ingresos mensuales, esta variable ha sido utilizada por otros autores dado que alcanza a representar el rendimiento de la persona por cada hora de trabajo; además, se define como indicador de salud a la variable dummy que indica si ha estado enfermo en las últimas 4 semanas.

Se encontró principalmente que el impacto general de la salud sobre la productividad es significativo y positivo (2.67%); además, respecto a los niveles educativos, el nivel superior posee la tasa de retorno de salud más alta (19.18%); respecto a los dominios geográficos, en Lima Metropolitana se halla el mayor impacto positivo y significativo de la salud (11.60%); respecto a los quintiles de ingreso, en quinto quintil se encuentra el mayor impacto (3.70%); por último, respecto a los rangos de edad, en el rango de 35-44 años, la salud tiene el mayor impacto sobre la productividad (5.35%). En general, vemos que el impacto de la salud sobre la productividad será mayor en tanto tengan un nivel educativo superior, reciban mayores ingresos mensuales y pertenezcan a las cohortes superiores de ingresos y edad.

Línea de investigación: 5300 – 2.b5

Palabras clave: Salud, productividad, modelo de Heckman, niveles educativos, dominios geográficos, quintiles de ingreso, rangos de edad, ingreso mensual e indicador de salud.



ABSTRACT

The main objective of this research work is to identify the impact of health on productivity in Peru for the year 2019, with respect to 4 disaggregations, these are: educational levels, geographic domains, income quintiles and age ranges. For this purpose, an Ordinary Least Squares (OLS) econometric model corrected by Heckman's two-stage estimation was used, which has the advantage of correcting the selection bias that occurs in the estimation of the wage equation. The study covers 22,061 respondents selected from the National Household Survey (ENAHU); For the analysis, the explanatory variables are grouped according to individual characteristics, human capital variables, housing variables and local market variables. Likewise, given that productivity cannot be measured in a traditional way, this research work establishes the variable of monthly income as a productivity indicator, this variable has been used by other authors since it reaches to represent the performance of the person by every hour of work; In addition, the health indicator is defined as the dummy variable that indicates whether he has been sick in the last 4 weeks.

It was mainly found that the general impact of health on productivity is significant and positive (2.67%); In addition, regarding educational levels, the higher level has the highest health return rate (19.18%); Regarding the geographical domains, Metropolitan Lima has the greatest positive and significant impact on health (11.60%); Regarding the income quintiles, the greatest impact is found in the fifth quintile (3.70%); Finally, regarding the age ranges, in the range of 35-44 years, health has the greatest impact on productivity (5.35%). In general, we see that the impact of health on productivity, will be greater if they have a higher educational level, receive higher monthly income and belong to the higher income and age cohorts.

Line of research: 5300 – 2.b5

Keywords: Health, productivity, Heckman model, educational level, geographic domains, income quintiles, age ranges, monthly income and health indicator.



INTRODUCCIÓN

El tema que se trata en el presente trabajo de investigación es el impacto de la salud en la productividad de individuos con residencia en Perú en el año 2019, añadiendo a su vez análisis desagregado por dominios geográficos, rangos de edad, niveles educativo y quintiles de ingreso. Para ello, tomaremos como muestra a los individuos encuestados en la Encuesta Nacional de Hogares sobre Condiciones de Vida y Pobreza en el año 2019.

Empezamos afirmando que resulta coherente pensar que las personas mejor alimentadas y con buena salud física y mental pueden tener un mejor rendimiento en las diferentes etapas de su vida. En un comienzo permite un mayor aprovechamiento de la enseñanza en la etapa estudiantil, luego, ese factor repercute en su desempeño laboral y, en última instancia, se refleja en sus ingresos. Esto último, influirá directamente en el presupuesto disponible del individuo para invertir en salud, por tanto, este proceso se convierte en un círculo virtuoso que permitirá generar una mayor riqueza para el individuo.

El tema del impacto de la salud en la productividad ha ido cobrando mayor relevancia en los últimos años; sin embargo, ha sido un tema poco tratado en el Perú. Rafael Cortez ha realizado los mayores aportes para el caso peruano con sus publicaciones en 1999 y 2002, él realizó un estudio desagregado del efecto de la salud sobre la productividad basándose en la zona rural/urbana y género.

El estudio que proponemos, dada la importancia de la variable salud sobre el impacto en la productividad, resulta interesante, pues interioriza más en un tema que no se había trabajado hace más de 18 años en el Perú, además, el nivel de desagregación es distinto a los anteriores planteados. En específico, planteamos una desagregación por dominios geográficos considerando que las realidades de las tres regiones (Costa, Sierra y Selva) son distintas para la zona norte, centro y sur. Asimismo, consideramos la desagregación por rangos de edad, quintiles de ingreso y niveles educativos, ya que son características particulares que pueden brindar un nuevo aporte significativo en el estudio del tema para el caso del Perú.

El interés del presente trabajo de investigación es, en primer lugar, determinar el impacto de las condiciones de salud en la productividad; y, en segundo lugar, realizar un análisis desagregado del impacto de la salud sobre la productividad dividido por dominios geográficos, rangos de edad, niveles educativos y quintiles de ingreso.

Algunos autores que han estudiado la relación entre la salud y la productividad como Cortez (1999) y Ribero y Núñez (1999), han esperado que una buena salud sea un indicativo de mayor productividad; sin embargo, han encontrado dificultades, principalmente en la medición de la salud. Esta medición resulta difícil por varias razones, Schultz y Tansel (1997), por ejemplo, manifiestan que hay varios indicadores que podrían obstaculizar la medición de la salud, como son las tasas de reportes de enfermedad, las tallas de adultos o, incluso, los días de enfermedad presentan problemas de errores en cuanto a su medición y no miden un rango completo de condiciones. A pesar de lo antes mencionado, la variable salud se podría establecer como los días de enfermedad, tal como lo han utilizado Cortez (1999) y Ribero y Núñez (1999), quienes ya estudiaron el tema y obtuvieron resultados, es decir, los días en que una persona se encontró inhabilitada de trabajar por algún tipo de enfermedad.

Por otro lado, la productividad estaría establecida a través de los salarios y, a partir de ello, un factor relevante que sobresale es el de determinar cómo la salud afecta a los salarios tanto para jóvenes y adultos según la región a la que pertenecen.

En ese sentido, en la presente investigación establecemos a la salud como uno de los determinantes del capital humano y por ende de la productividad, lo cual permite determinar el impacto sobre diversas características representativas de los individuos de la muestra como: el dominio geográfico en que se encuentran, el rango de edad al que pertenecen, el quintil de ingreso en que se hallan y el nivel de educación que tienen. Primero, con respecto al dominio geográfico, esperamos encontrar diferencias claras del impacto de estudio entre los ocho dominios geográficos que se tratarán en la presente investigación, en la que esperamos que la diferencia sea mayor en zonas de sierra y selva en comparación con la zona de costa y Lima Metropolitana, esto se podría plantear inicialmente debido a la mayor necesidad de capacidad física en los trabajos de los individuos de los primeros sectores mencionados, por el menor desarrollo de actividades industrializadas como en la costa y Lima.

Segundo, en cuanto a los rangos de edad, el trabajo planea sustentar por qué la diferencia del impacto de la salud sobre la productividad sería mayor en los rangos de edad más jóvenes, nosotros anticipamos que se debería a que los individuos que pertenezcan a rangos de edad más jóvenes se dedicarían a labores que tomen mayor esfuerzo físico que los que pertenecen a rangos de edad adultos.

Tercero, en cuanto a los quintiles de ingreso, esperamos que la diferencia sea mayor para los individuos pertenecientes a los quintiles más pobres debido a que contarían con un menor presupuesto para la inversión en su salud.

Finalmente, el impacto de la salud sobre la productividad desagregado en los tres niveles educativos básicos mostraría la mayor diferencia en el nivel educativo primaria que en el nivel secundaria o superior, esto lo esperaríamos debido a que individuos con menor nivel educativo se dedicarían a labores de mayor esfuerzo físico y poseer buena salud sería fundamental para realizar las actividades que determinarían su salario.

El presente trabajo constituye una mirada innovadora sobre al tema de salud y productividad, ya que incorpora una desagregación diferenciada de trabajos de casi 20 años de antigüedad, y así sentamos un estudio significativo para el caso del Perú. Es por ello que, establecemos objetivos, así como preguntas que respondemos mediante hipótesis acerca de lo que esperamos encontrar en el presente trabajo de investigación, las cuales se pasarán a explicar a continuación.

Iniciamos estableciendo el objetivo general del presente trabajo de investigación, que es determinar el impacto de la salud sobre la productividad en el Perú para el periodo del año 2019. A partir del objetivo general, establecemos cuatro objetivos específicos para profundizar en el tema. El primero es determinar la diferencia del impacto de la salud sobre la productividad entre los ocho dominios geográficos: Costa Norte, Costa Centro, Costa Sur, Sierra Norte, Sierra Centro, Sierra Sur, Selva y Lima Metropolitana. El segundo es determinar la diferencia del impacto de la salud sobre la productividad entre cinco rangos de edad: 18-24, 25-34, 35-44, 45-59 y 60-70. El tercero es determinar la diferencia del impacto de la salud sobre la productividad entre los quintiles de ingreso. Finalmente, el cuarto objetivo específico es determinar la diferencia del impacto de la salud sobre la productividad entre los tres niveles educativos: primaria, secundaria y superior.

La pregunta general es: ¿Cuál es el impacto de la salud sobre la productividad en el Perú?, y la hipótesis general para ella es que el impacto de la salud sobre la productividad en el Perú es positivo y significativo durante el periodo de estudio.

Las preguntas específicas e hipótesis que planteamos son cuatro. La primera es: ¿En qué medida se diferencia el impacto de la salud sobre la productividad entre los ocho dominios geográficos (Costa Norte, Costa Centro, Costa Sur, Sierra Norte, Sierra Centro, Sierra Sur, Selva y Lima Metropolitana)?, y la hipótesis específica con respecto a esa pregunta es que el impacto de la salud sobre la productividad, entre los ocho dominios geográficos, es mayor para los dominios geográficos situados en la zona sierra y selva, en comparación con la zona costa y Lima Metropolitana. Lo anterior se debe al menor desarrollo de las actividades industrializadas en las zonas de la sierra y selva, que requieren personas con mayor capacidad física para realizar trabajos relacionados con la manufactura, agricultura, entre otras; en ese sentido, el impacto de la salud sobre la productividad será mayor en estas regiones donde se requiere principalmente de mayor esfuerzo físico para realizar sus labores.

La segunda pregunta específica es: ¿En qué medida se diferencia el impacto de la salud sobre la productividad entre los cinco rangos de edad (18-24, 25-34, 35-44, 45-59, 60-70)?, y la hipótesis a la pregunta mencionada es que el impacto de la salud sobre la productividad, entre los cinco rangos de edad, es mayor para los rangos de edad más jóvenes, ya que generalmente se dedican a labores que demandan mayor esfuerzo físico, en comparación con los más adultos que generalmente se dedican a labores que demandan, principalmente, más experiencia.

La tercera pregunta específica es: ¿En qué medida se diferencia el impacto de la salud sobre la productividad entre los quintiles de ingreso?, y la hipótesis que esperamos para esta pregunta es que el impacto de la salud sobre la productividad, entre los quintiles de ingreso, es mayor para los quintiles más pobres, ya que cuentan con un menor presupuesto para invertir en su salud, por tanto, están más propensos a estar enfermos y padecer más dolencias que afecten su productividad.

La cuarta y última pregunta específica es: ¿En qué medida se diferencia el impacto de la salud sobre la productividad entre los tres niveles educativos (primaria, secundaria y superior)?, y la hipótesis respectiva es que el impacto de la salud sobre la productividad, entre los tres niveles educativos, es mayor para los niveles educativos más básicos, en

ese sentido, habrá un mayor impacto en el nivel primaria, en comparación con el nivel secundaria y superior; y, habrá un mayor impacto en el nivel secundaria, en comparación con el nivel superior. Lo anterior se debe a que las personas con un menor nivel educativo se dedican a labores que generalmente implican mayor esfuerzo físico, por tanto, mantenerse con buena salud debe ser primordial, ya que de eso depende su trabajo y su salario.

Cabe resaltar que la presente investigación pertenece a la línea y sub-línea de investigación “Consecuencias económicas de la falta de salud” dentro del área de Economía de la Salud del Instituto de Investigación Científica (IDIC) de la Universidad de Lima, cuyo código es 5300 - 2.b5.



CAPÍTULO I: ANTECEDENTES

La salud es un requerimiento de los seres humanos para obtener una vida duradera, además, es un determinante de la productividad del capital humano, ya que permite un correcto aprovechamiento de las actividades de una persona en las diferentes etapas de su vida. Los primeros estudios teóricos y empíricos establecían que la inversión en capital humano se daba exclusivamente a partir de la educación. Comenzando la década de 1970 varios autores empezaron a considerar como determinantes del crecimiento económico: la inversión en salud y la expectativa de vida; y a partir de ello han planteado modelos desde un escenario de generaciones traslapadas.

Un primer precedente trascendente para la consideración de la salud como un determinante del crecimiento económico es el que realizan Ehrlich y Lui (1991), el cual construye un modelo de generaciones superpuestas que destaca la importancia de la institución familiar y la importancia de capital humano como motores del crecimiento económico de un país. Ehrlich y Lui (1991) explican que este modelo se basa en que los padres invierten en sus hijos con la finalidad de que estos cuiden de ellos en su vejez, maximizando las oportunidades de crecimiento a través de un proceso de optimización intertemporal. El rol de la salud en este proceso se da en las probabilidades de sobrevivencia en la adultez mayor, midiendo así la longevidad, que si es mayor da lugar a un incremento de la inversión del capital humano y, por ende, un aumento también en la tasa de crecimiento a largo plazo.

El Perú, desde finales de los años sesenta, en el contexto de una época alejada de la democracia, hasta la primera década del siglo XXI, se mantuvo alejado del desarrollo en salud que experimentaron diversas naciones de América del Sur. La salud peruana fue un ámbito muy poco desarrollado que, a lo largo del nuevo milenio, tuvo grandes oportunidades de surgimiento con el Acuerdo Nacional de julio de 2002 y la Reforma de Salud del 2013; sin embargo, solo sirvieron como medios para relucir el ineficiente sistema de esos momentos y la poca valoración a un recurso tan importante como la salud.

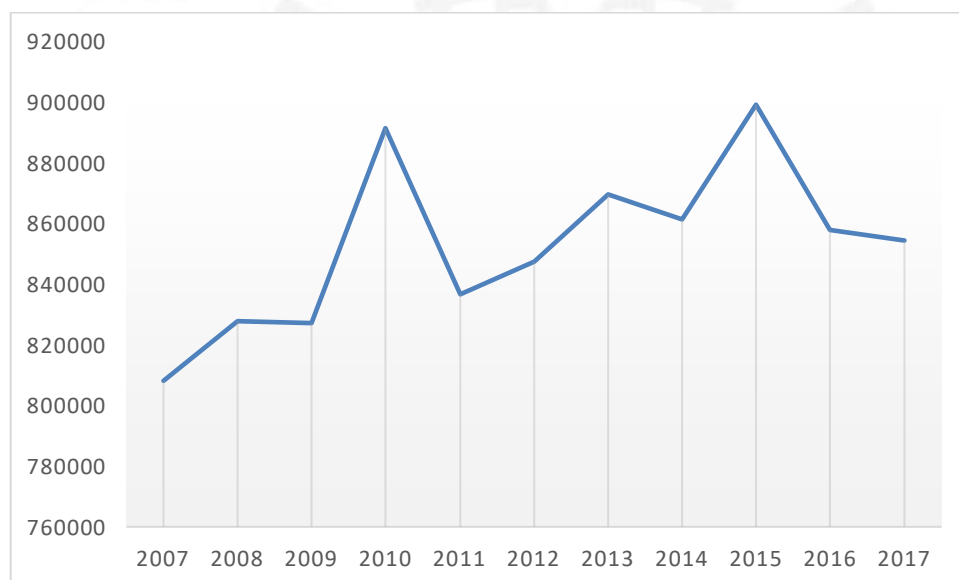
Pese a todo ello, cabe resaltar que el país cuenta con un Seguro Integral de Salud (SIS) con alto nivel presupuestario, el cual es considerado por la Organización Mundial

de la Salud (OMS) como uno de los mejores debido a la amplia cobertura que brinda el seguro sin ningún condicionamiento político o interés en el medio. Con respecto a esto último, sobresale un objetivo que debe prevalecer en una nación, esto es, el derecho fundamental al cuidado integral de la salud, el cual se podría ver reflejado en la mayor productividad de los individuos del país y promoviendo un mayor crecimiento económico.

La evolución del número de hospitalizados del año 2007 al 2017, como se visualiza en la Figura 1.1, ha mostrado un significativo aumento que puede ser una muestra a grandes rasgos de cómo se encuentra la salud en el Perú, medida a través del número de personas que se encuentran mal de salud y son hospitalizadas.

Figura 1.1

Evolución del número de hospitalizados en Perú desde el año 2007 al 2017



Fuente: INEI.

Elaboración Propia.

Desde una perspectiva aún más específica, se puede definir a la variable salud como el número de días que se encuentran enfermos los individuos, como lo han realizado autores como Cortez (1999) y Ribero (2000) en sus investigaciones con el fin de relacionarla con la productividad. Determinar ese vínculo es un tema relevante para el estudio, ya que permitiría visualizar el impacto del periodo de incapacidad, originado por la enfermedad del individuo, sobre la productividad, no solo para beneficio de estos, sino,

también para el del crecimiento económico del país a través del aumento en la productividad.

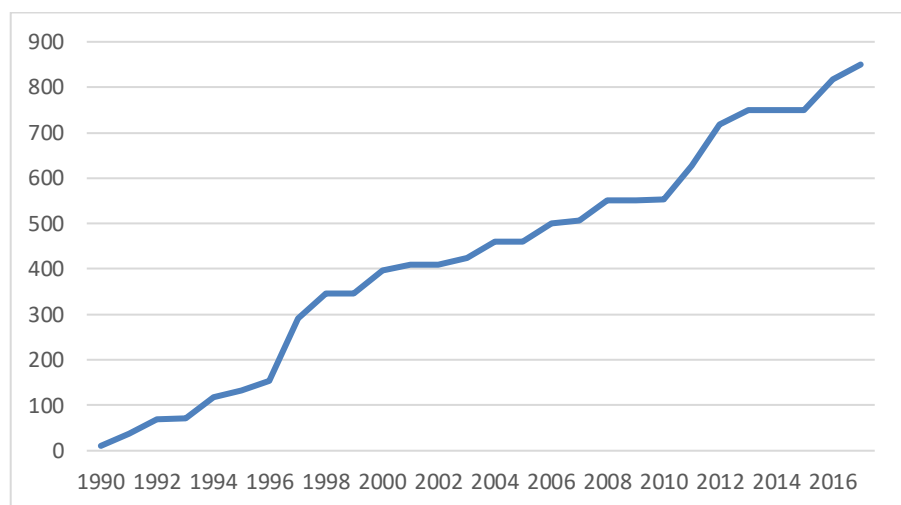
Según el INEI, en la actualidad, el 34.8% del salario mínimo está designado a cubrir la canasta básica familiar en un hogar conformado por 4 personas. La canasta básica, según el Ministerio de Salud, es un conjunto de servicios y productos de primera necesidad que una familia necesita para subsistir durante un mes. Los productos están cubiertos en su mayoría por alimentos; y en cuanto a los servicios se incluyen el de transporte, higiene, vestuario y salud. Es por ello, que la evolución del salario mínimo a lo largo de los años establece también en cierto modo la evolución de la canasta básica familiar. La importancia de corroborar que una familia con salario mínimo cubre con la canasta básica es muy importante para el Estado, ya que a partir de ello determina la revisión del salario mínimo y la vigilancia de los precios de alimentos básicos, así como el acceso a servicios para la salud.

La Figura 1.2, como ya veníamos mencionando, muestra la evolución de la remuneración mínima vital desde el año 1990 al 2016 y a partir de ello percibimos un incremento a lo largo de los años. El aumento de este ha sido un tema muy polémico debido a que ese incremento, según bases de grandes empresarios, atenta al crecimiento bajo el supuesto de economías decrecientes; sin embargo, el aumento de este se da en la misma proporción en cuanto al incremento del valor de la canasta básica familiar.

Es por ello que, en específico, el servicio de salud que está incluido dentro de la canasta básica, y, el gasto de este último dentro del salario mínimo vital puede confirmar en un primer momento que los individuos que cuentan con acceso a la salud proveniente de la canasta básica no afectan su productividad, pues cuentan con ese servicio. Esta última afirmación suena muy apresurada basándonos en un solo gráfico acerca de la evolución del salario mínimo vital y la inclusión de la canasta básica como parte de ella, por lo que el presente trabajo buscará determinar con mayor contundencia si es que un impacto en la salud afecta la productividad en los individuos del Perú en un contexto actual como el del año 2019, considerando más variables que solo el salario.

Figura 1.2

Evolución de la remuneración mínima vital en Perú desde 1990 hasta el 2017



Fuente: INEI

Elaboración Propia

Estudios previos que tratan el impacto de la salud sobre la productividad, como el de Cortez (1999), utilizan la variable ingreso mensual de los individuos como dependiente para el estudio del impacto de la salud en la productividad. Como se puede visualizar en la Figura 1.3, el ingreso promedio mensual lo lidera la región de la Costa por encima de la Sierra y la Selva, esta variable es muy importante dentro de la productividad, pues muestra en parte cómo estaría avanzando los salarios por cada región del Perú. Sería relevante, entonces, analizar cómo los niveles de ingreso establecidos a través de quintiles de ingreso son afectados por un impacto en la salud en las regiones del país, además de poder si un impulso en mejoras de salud en las regiones de la Sierra y Selva les permitirían incrementar su productividad y acercarse cada vez más a una convergencia en cuanto a la productividad con respecto a Lima.

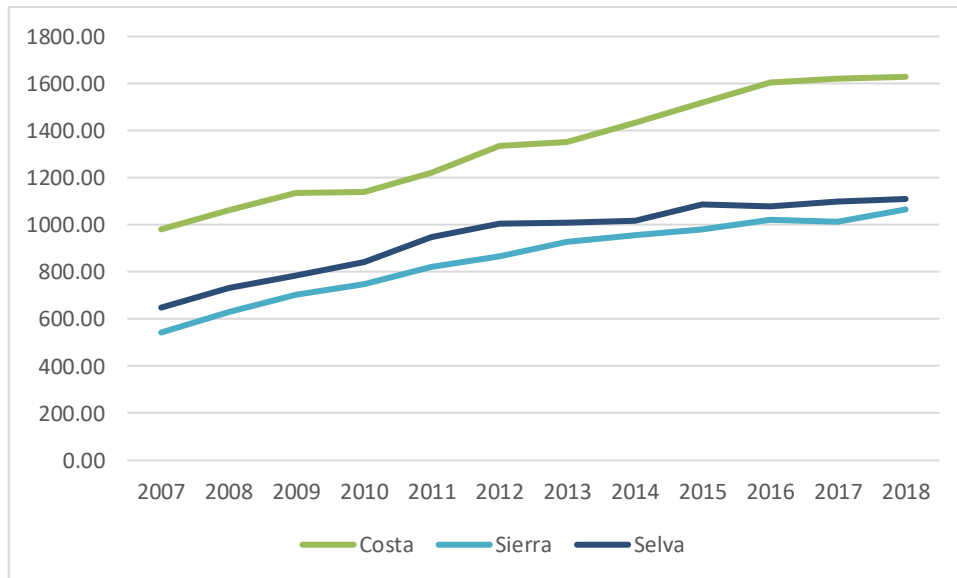
La esperanza de vida promedio actualmente en Perú es de 73 años, en comparación al periodo de 1950 – 1955 donde la esperanza de vida era de 44 años. Este aumento muy significativo se dio en gran parte al mejoramiento de accesos a servicios de saneamiento básico, cambios en los estilos de vida, al proceso de urbanización y a la extensión de la cobertura de servicios de salud. Este último factor es muy importante, ya que determina cuantas personas están incluidas en un seguro de salud.

En el Perú, el seguro social cuenta con un financiamiento superior a otros rubros de ámbito social y es considerado por diversas naciones europeas como el mejor sistema

nacional de salud por predominar el derecho fundamental de la salud sobre intereses netamente económicos.

Figura 1.3

Evolución del ingreso promedio mensual como indicador de la productividad proveniente del trabajo por regiones del Perú desde 2007 hasta el 2018



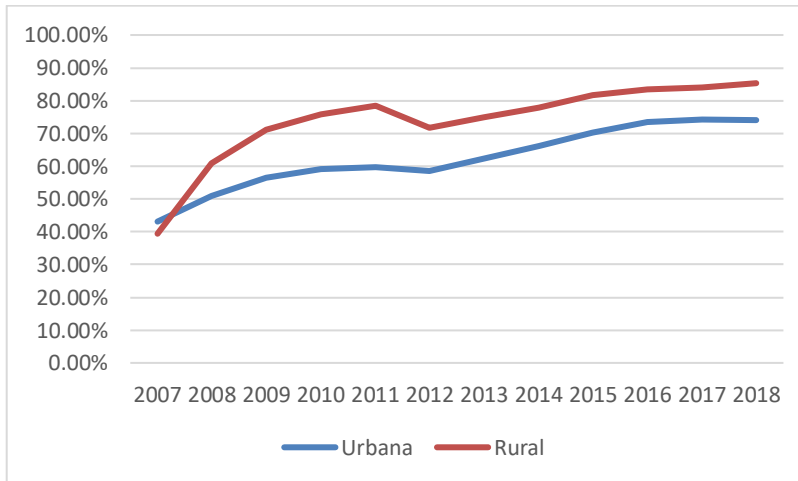
Fuente: INEI
Elaboración Propia

En el Perú, el seguro social cuenta con un financiamiento superior a otros rubros de ámbito social y es considerado por diversas naciones europeas como el mejor sistema nacional de salud por predominar el derecho fundamental de la salud sobre intereses netamente económicos.

En la Figura 1.4 se muestra el aumento de la población afiliada a un seguro de salud entre los años 2007 al 2018 según la zona de residencia. El sector público ofrece a la población en condiciones de pobreza y pobreza extrema el Seguro Integral de Salud (SIS) e independientemente de ello, el sistema de seguridad social incluye el seguro social con provisión tradicional a través de ESSALUD y la provisión privada a través de la EPS. En términos generales, la población tiene mayor capacidad de estar afiliada a un determinado seguro de salud según el contexto y circunstancias en que se encuentre. A raíz de ello, resalta la importancia de un estudio que determine el impacto del acceso a los servicios de salud sobre el nivel de salud reportado, y el impacto derivado de este último sobre la productividad del individuo..

Figura 1.4

Evolución de la población afiliada a seguro de salud según zona en Perú desde 2007 hasta el 2018

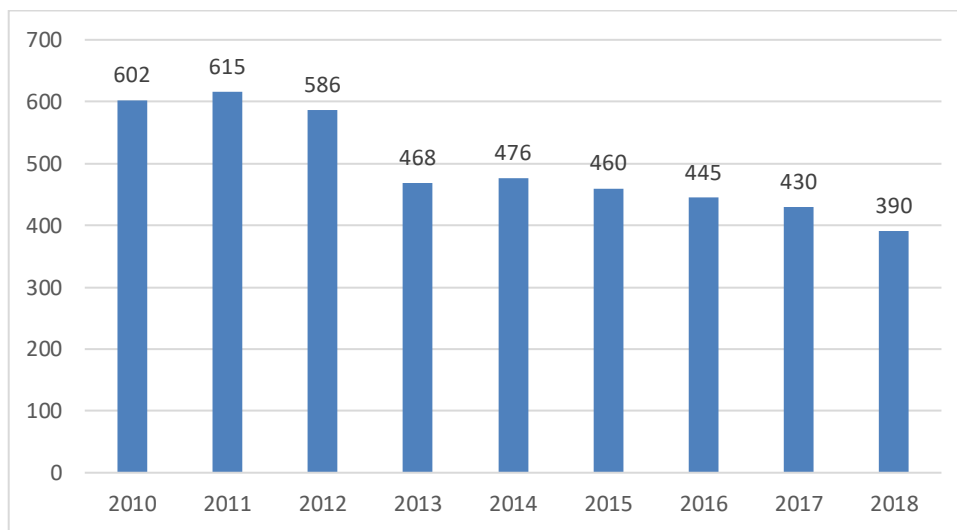


Fuente: INEI
Elaboración Propia

Como se puede apreciar en la Figura 1.5, desde el 2010 hasta el 2018 hubo una creciente oferta de especialistas en recursos humanos de la salud, disminuyendo así la ratio de habitantes por cada médico. Eso podría significar un adecuado equilibrio entre oferta y demanda y más aún si se trata de algo tan importante como la salud; sin embargo, también podría significar un riesgo en la calidad de la entrega de servicios de salud al incrementarse el número de especialistas con una formación académica menos rigurosa. En ese caso, resultaría adecuado medir la calidad del servicio de salud con el número de días enfermo que se encuentra un individuo. En el presente trabajo se realizará aquello para medir el impacto sobre la productividad.

Figura 1.5

Número de habitantes por cada médico en Perú desde 2010 hasta el 2018



Fuente: INEI

Elaboración Propia



CAPÍTULO II: REVISIÓN DE LITERATURA

2.1 Marco teórico

La relación entre la salud y la productividad han sido estudiadas bajo la teoría del capital humano, donde se argumenta que la educación y la salud determinan una mayor productividad y crecimiento económico. Esta relación se extiende desde los albores de la economía, cuando Adam Smith (1776) expresa que el hombre que se ha instruido con trabajo e invirtiendo tiempo para ello, debe estar ocupado en una labor que le permita un retorno no menor al costo en el que incurrió para educarse; es decir, debería percibir los beneficios de un capital comparable en valor de su formación. Largo tiempo transcurrió hasta que a mediados del siglo XX algunos autores como Schultz (1961), Becker (1964) y Mincer (1974) realizaron aportes a la teoría del capital humano

Schultz (1961) da el nombre a la teoría del capital humano en su artículo *Investment in Human Capital* (1961), donde plantea que el conocimiento junto con el desarrollo de aptitudes y habilidades son factores que determinan el nivel del bienestar. Además, Schultz muestra que el capital humano agrupa componentes cualitativos como los conocimientos, la habilidad y otros atributos que impactan sobre la capacidad para realizar un trabajo productivo; por último, menciona que las inversiones realizadas para desarrollar más estas capacidades tienen un efecto positivo sobre la productividad y tendrán un retorno positivo.

Tiempo después, Becker (1964) desarrolló formalmente los aportes de Schultz, Becker formula que el ámbito educativo es el factor principal para el capital humano, por ello, cuando hay mayor inversión en el sector educativo, vía gasto público, se eleva el ingreso de los individuos por el nexo que existe entre el crecimiento económico y el stock de capital humano. Entonces, Becker (1964) en esencia planteó que la educación y la formación, que comprenden conocimientos generales y específicos, son inversiones que los individuos realizan para aumentar su productividad y finalmente sus ingresos;

además, supone que el individuo que decidirá sobre invertir o no en educación, comparará los costos y beneficios que obtendrá en el futuro si decide educarse más.

Luego, Mincer (1974) incluyó a la teoría del capital humano, la escolaridad y la experiencia como variables que impactaban positivamente en los ingresos de los individuos. Mincer realizó dos aportes principales al concepto del capital humano, ambos los realizó bajo la guía y enseñanzas de Becker; el primer aporte está relacionado con la fórmula del valor presente para calcular los retornos del entrenamiento, esta se muestra a continuación:

$$\frac{d}{c} = (1 + r)^n$$

Donde:

r: tasa de retorno de la inversión

c: costo de entrenamiento

d: incremento de las ganancias luego del entrenamiento

n: número de años del entrenamiento

En ese primer aporte, Mincer plantea que las ganancias incrementales producto de los años de estudios son directamente proporcionales con las ganancias que se dejan de recibir durante el entrenamiento y, también, varía directamente proporcional con la tasa de retorno y el número de años de instrucción. Luego de este primer aporte, Mincer realiza su segundo y mayor aporte a la teoría del capital humano, en este segundo aporte Mincer (1974) propone una regresión lineal como metodología para medir el impacto de la escolaridad y la experiencia sobre los ingresos de los individuos. La regresión de Mincer incorpora al logaritmo del ingreso como variable dependiente, mientras que las variables explicativas son la escolaridad y los años de experiencia. La función propuesta por Mincer se muestra a continuación:

$$\ln Y = \beta_0 + \beta_1 S + \beta_2 X + \beta_3 X^2 + \varepsilon$$

Donde Y son los ingresos para un periodo definido, S representa los años de educación, X es la experiencia y X² es la experiencia al cuadrado que representa la no linealidad que existe entre la edad y los ingresos; además, teóricamente se concluye que los coeficientes de la educación y la experiencia deben ser positivos, mientras que el

coeficiente de la experiencia al cuadrado debe ser negativo. Este aporte principal permitió que se realizarán mayores investigaciones en el campo de la teoría del capital humano, entre estas investigaciones se interioriza la relación de la productividad con el factor de la salud, la cual es de mayor relevancia para el estudio y son las que se mencionarán en las siguientes líneas.

Ahora bien, estos estudios anteriores limitaban el análisis del capital humano a la educación y la experiencia como factores para el incremento de la productividad. A finales de la década de 1980, se formularon los primeros estudios que vincularon la salud con la productividad, esta relación se evidenció en el marco de la hipótesis del salario de eficiencia (Pitt, Rosenzweig y Hassan, 1990; Behrman y Deolalikar, 1988; Sahn y Alderman, 1988), desde ese momento se realizaron mayores aportes que asociaban la productividad y la salud.

Al trabajo de Schultz de 1961 que planteó la inversión en salud como variable determinante del capital humano, le siguió el trabajo de Lewis (1980) quien planteó directamente a la alimentación y la salud como factores determinantes de la productividad, por tanto, sugería que las empresas debían asegurar a sus trabajadores con: servicios de salud adecuados, dietas alimenticias apropiadas y un lugar de reposo tranquilo. Lewis afirma que estos factores garantizarían un incremento de la productividad de los trabajadores y, por tanto, mayores beneficios para el empresario. Este aporte de Lewis es coherente con el que expuso Mushkin (1962) que establecía que las inversiones en servicios de salud y educación aumentarían la productividad de los trabajadores, específicamente sugirió que las políticas enfocadas al sector salud impactarían positivamente sobre el producto y generaría un rendimiento en el largo plazo.

Un aporte preciso que hace Schultz (1983) es sobre cinco factores que ayudan al mejoramiento del capital humano, estos son: los servicios de salud con su equipamiento correspondiente, la formación en el puesto de labores, la educación formal que se obtiene en el nivel primaria, secundaria y superior, los programas de extensión y de estudio para los adultos; y la emigración de agentes económicos para adaptarse al contexto cambiante de oportunidades laborales.

Con lo antes mencionado, Schultz (1983) propone explícitamente brindar mayor relevancia a factores relacionados con el aprendizaje y la salud de los trabajadores para mejorar su productividad.

Ahora bien, el modelo principal acerca del capital humano con relación a la salud es desarrollado por Grossman (1972) que plantea principalmente que la productividad de un individuo estaría representada por el tiempo que puede pasar el individuo generando ingresos y riqueza, todo lo anterior bajo un enfoque de salud. La relación que muestra es que mientras el individuo tenga un mejor estado de salud, entonces, el tiempo perdido por estar enfermo y sufrir dolencias será menor, lo cual afectará positivamente sobre su productividad. En ese sentido, habrá una mayor demanda de los trabajadores por salud, ya que al incrementarse este factor, entonces, habrá un efecto positivo sobre su productividad.

A modo de conclusión, sobre lo propuesto por Grossman (1972), es que propone una alternativa para valorar la salud de los individuos, ya que generalmente los datos proporcionados por los individuos están sujetos a errores de medición. Esta alternativa que propone es la cantidad de tiempo sin estar enfermo (sano), esto se define como el stock de salud; por lo tanto, con esta alternativa se puede inferir el estado de la salud del individuo que fue parte de la muestra y poder realizar las regresiones correspondientes en los trabajos que contengan un apartado empírico.

2.2 Literatura empírica

Esta sección contiene los hallazgos empíricos en diversos países. En general, los estudios de salud y productividad han llegado a consensos muy similares del efecto positivo de la salud sobre la productividad. La medición de las condiciones de salud por parte de la literatura económica se ha enfatizado en indicadores recogidos por las Encuestas de Hogares, entre los cuales destacan, por ejemplo: talla de adultos, los días de enfermedad y los días de incapacidad. Dichas variables no están libres de error de medición debido a sesgos que están correlacionados con el nivel de educación de la persona a la que se dirige, los ingresos de su hogar y otras variables no observadas.

La mayoría de los estudios que se presentan a continuación se dan en el contexto de Latinoamérica y a partir de los años 90. Uno de los primeros precedentes así es el que se realiza en Colombia, en donde Fogel (1994) confirma que la estatura, variable que utiliza para reflejar la salud, está relacionada positiva y significativamente con los salarios

de los individuos de su muestra. Encuentra, además, un vínculo de la estatura e índice de masa muscular con la mortalidad de hombres mayores. Unos años más tarde, con el mismo propósito de explicar el impacto de la salud sobre la productividad, Thomas y Strauss (1997) utilizaron diferentes dimensiones de la salud, como la información de la talla de adultos y el índice de masa muscular (IMC) para ver cómo afectaban la productividad a través de los salarios de hombres y mujeres en las zonas urbanas de Brasil. Las principales conclusiones de los autores a partir de ello fueron que la productividad es explicada a través de un efecto positivo y significativo de la talla de los individuos y el IMC, además, también concordaron que un incremento de IMC en hombres no educados se asociaba con salarios aún más altos que hombres educados con IMC alto.

Un trabajo con mayor énfasis en explicar a la variable salud en ese mismo año fue el que realizaron Schultz y Tansel (1997) que explicaron a la variable salud a través de los días de inhabilitación con el objetivo de obtener ecuaciones estimadas de los salarios de Costa de Marfil y Ghana. Los resultados que esperaron con ello era que las personas sean menos productivas al estar enfermas; sin embargo, en su estudio encontraron problemas para comprobar ese efecto como la medición de la morbilidad en los adultos en las encuestas de hogares, así como el establecer el impacto de la morbilidad en la productividad. Su principal hallazgo fue determinar cómo el nivel de salarios y una larga vida productiva son explicados por parte de las condiciones de salud, por otro lado, encontraron que la solución a los problemas de comprobación era la utilización de variables instrumentales como los precios de los alimentos locales y de los servicios de salud. En el presente trabajo de investigación, tomando en cuenta la solución a los problemas de medición que plantean estos autores, también utilizaremos variables instrumentales como el acceso a la electricidad, a la red telefónica, entre otros que serán explicados más adelante en la metodología que se utilizará.

Asimismo, el año de 1999 traerá consigo más de un trabajo de investigación con respecto al tema de impacto de la salud sobre la productividad, uno en específico en el cual el presente trabajo de investigación tomará mucha inspiración es el de Cortez (1999) en el cual se estudia el efecto de la salud, medido a través del número de días de enfermedad, sobre el salario por hora de mujeres y hombres de edad adulta, según el área urbana o rural, en Perú con la finalidad de determinar la productividad en torno a la salud.

Sus datos fueron obtenidos de la Encuesta Nacional de Hogares (ENAH) de 1995 con una delimitación de 19.975 hogares del país y su principal conclusión es que, a partir de la estimación de un modelo MCO corregido por la estimación en dos etapas de Heckman, la productividad es explicada por parte de la salud. Sugiere también que la inversión en salud, en el ámbito privado como público, debería ser un factor decisivo para el aumento del bienestar de los hogares a través de sus ingresos.

Otros trabajos de relevantes de ese año fueron los que se realizaron en México, por ejemplo, en el que Parker (1999) realizó un estudio en personas adultas mayores de dicho país utilizando la Encuesta nacional de edades de México con la finalidad de determinar los factores que determinaban su salud. Demostró que un nivel de salud bajo disminuía en un 58% el salario de las personas mayores y que tanto la vejez como la pobreza están estrechamente relacionadas. Su principal conclusión fue que condiciones deficientes de salud inciden en bajos ingresos, y a partir de ese vínculo sugiere una correcta aplicación de políticas de salud con la finalidad de combatir la pobreza en personas jubiladas. Además, Knaul (1999) utilizó la Encuesta Nacional de Población del año 1995 de México con la finalidad de evaluar el efecto de inversión en salud y nutrición. en edad temprana. sobre los salarios a largo plazo, utilizando como principal indicador la edad de la primera menstruación. Utiliza este indicador debido a que el inicio de este ciclo en las mujeres resume un conjunto de complejos procesos de desarrollo en el sistema nervioso y central y sistemas endocrinos que conducen a aumentos de la fuerza y la resistencia, que a su vez se demuestra en la realización de sus actividades laborales. Sus resultados arrojaron que, a una disminución de los años de ocurrencia de la primera menstruación, el nivel del salario aumenta por hora entre un rango de 23% y 26%. Por último, concluye que el acceso a los establecimientos de salud tiene un efecto marginal y no relevante.

También en Perú, Murrugarra y Valdivia (1999) identificaron cómo los determinantes del estado de salud de adultos de zonas urbanas del Perú afectaban a sus salarios. Para ello utilizaron datos de la Encuesta de niveles de vida (ENNIV) del año 1994 en un modelo multinomial logit. Sus principales resultados demostraron que los efectos de una salud adecuada sobre la productividad son menores en trabajadores formales que en personas autoempleadas, los primeros reflejan una disminución en sus salarios de 1.2%, mientras que los segundos presentan una disminución en 3.2% por día

de enfermedad adicional. Otra conclusión del estudio fue encontrar que las personas mayores enferman con mayor frecuencia y en un rango de tiempo más largo, además, encuentran que la salud se ve afectada de manera diferenciada, según la edad, por los años de educación de los individuos.

En Colombia, Ribero y Núñez (1999) descubrieron que no encontrarse bien de salud en el transcurso de un mes disminuye en 14% y 28% los salarios en mujeres y hombres de la zona rural, respectivamente. Además, establecieron una relación en cuanto a la talla y los salarios y descubrieron que un centímetro adicional aumenta en los salarios en 7% y 8% para mujeres y hombres, respectivamente.

Ya en la década de los 2000 hacia adelante se presentan más trabajos de investigación con respecto al tema con mayor precisión en la definición de variables como Ribero (2000) en el que mide el efecto de la salud, a través de indicadores básicos, sobre la productividad laboral en Colombia, además, en su estudio sugiere que para que la productividad de los individuos mejore, el estado debe intervenir aplicando políticas públicas más especializadas en el sector salud. Los datos obtenidos para la investigación fueron extraídos de la Encuesta Nacional de Hogares de diciembre de 1991 y de la Encuesta de Caracterización Socioeconómica de 1993. Ribero (2000) estima en una función de ingresos de tipo Mincer de los indicadores básicos, bajo la suposición de que la variable salud es endógena, y encuentra a partir de ello que la salud cuenta con un signo acertado y significativo. La principal conclusión a la que llega es que un aumento del 50% en la media de días de inhabilitación se derivará en una disminución en los ingresos individuales, tanto para hombres y mujeres de zonas rurales y urbanas.

Picado y Vargas (2001) utilizan la base de datos de Encuesta de Hogares y Propósitos Múltiples de 1998 (EHPM) para extraer información de 43 944 individuos entre 12 y 70 años de edad de Costa Rica con la finalidad de medir el impacto de la salud sobre la productividad de estos a través de sus salarios. Asimismo, bajo el mismo criterio de diversos autores, Picado y Vargas (2001) emplean variables instrumentales para medir el impacto de la salud sobre la productividad en Costa Rica, es así que definen como indicadores de salud, dentro de su trabajo, a las variables DISCAP e INDICESALUD. La primera, una variable dummy que informa que es 1 cuando el individuo presenta algún tipo de discapacidad permanente, y 0 cuando no. La segunda, un indicador de 90 posibles estados de salud construidos con base en el número de hospitalizaciones y consultas

médicas recibidas por los encuestados. Luego, sobre la base de la construcción de la ecuación salarial y la regresión estadística, determinarán que el rol de salud tiene un impacto positivo en la productividad. La sugerencia que brindan a los gobiernos de países en desarrollo y organismos internacionales con respecto al rol de la salud como forma de capital humano, es la de impulsar una inversión adicional con la que supuestamente ya cuentan para un beneficio a largo plazo en mejoras de la productividad.

Ribera (2003) estudia el papel de la salud en la productividad en economías desarrolladas de 24 países pertenecientes a la OCDE. El objetivo del estudio es analizar la relación antes mencionada a través de una mejora en los niveles de capital humano. Los resultados que obtiene a través de una estimación del modelo SOLOW ampliado con SALUD es que la salud afecta de forma positiva el crecimiento económico. Además de ello, encuentra que las economías desarrolladas cuentan con variables explicativas como gasto sanitario, gasto corriente y gasto de capital. De todas esas variables, la primera es la única que tiene un efecto directo positivo en cuanto al aumento de la productividad, esto según la autora debería ser muy relevante en cuanto a la realización de políticas públicas.

Chakraborty y Das (2005) juntan las diferencias en el estado de salud entre pobres y ricos con su situación económica mediante la proposición de un mecanismo. A partir de ello resaltan que la productividad laboral de un individuo es mejorada a través de la educación y la inversión en salud de ámbito privado. Los autores también destacan el rol fundamental del capital de la salud en el camino al desarrollo, pues cuenta con un papel diferenciado al capital humano. La diferenciación de la salud que proponen recae en la capacidad de incentivar a los individuos a través de la definición del riesgo de mortalidad, ello nos indica que los individuos estarán más incentivados a trabajar desde más jóvenes cuando sepan su tiempo aproximado de vida.

Ahora, con respecto a estudios más recientes en el contexto de Latinoamérica, destaca el realizado por Rivera, Currais y Rungo (2008) extrajeron datos de la encuesta brasileña de hábitos para definir sus variables y fijar una relación intergeneracional entre ellas, todo ello con el objetivo de destacar el rol fundamental de la salud sobre la educación, el capital humano y la productividad transmisible de generación en generación. Las conclusiones a la que llegan es que un efecto significativo y positivo en la salud contribuyen al traslado intergeneracional de ingresos, en ese sentido, los hijos de

padres ricos tienen una mayor tendencia a la salud y a ser más productivos, además de acumular mayor capital humano; mientras que, para el caso de hijos de padres pobres presentan una tendencia de mantenerse en la pobreza y no salir de ese círculo vicioso.

Gupta y Barman (2010) intentan desarrollar un modelo de crecimiento endógeno resaltando el papel del capital de salud, infraestructura pública y contaminación. Es una extensión del modelo de Agenor (2008). Ambos concluyen que el ámbito de la salud debería de ser más especializado en cuanto a las políticas públicas que se aplican, con el objetivo de promover una mayor productividad y crecimiento económico.

Salas (2010) establece que la salud es una variable promotora de la productividad en las nuevas formas de organización laboral. Afirma que las empresas, antes de aplicar dichas medidas, buscaban ser más competitivas y eficientes, además de disminuir costos y realizar sus actividades en el menor tiempo posible a costa de que sus trabajadores estén sometidos a un exceso de esfuerzo que generaba un deterioro en la salud de ellos tanto físico como mental. Esto último ha provocado un perjuicio tanto para los empleados como para los empleadores, en los primeros en cuanto a perjudicar su salud y en los segundos en cuanto a una disminución del desempeño de sus trabajadores, generando todo lo contrario a lo que esperaban inicialmente. Es a partir de lo antes explicado que el autor resalta y sugiere el establecimiento de políticas de promoción y prevención de la salud dentro de la dirección y visión estratégica de todas las organizaciones y empresas que cuentan con un personal que labora bajo grandes exigencias. Destacamos que esta recomendación puede ser fundamental para todo tipo de trabajo de investigación que toque el tema del impacto de la salud en la productividad a un nivel micro para generar un beneficio absoluto en una compañía o empresa.

En la última década, han salido a relucir trabajos de investigación que miden el impacto económico de un estado de la salud como la desnutrición crónica y aguda en el contexto de Perú, ejemplo de ello es el estudio realizado por Alcázar et al. (2013), en donde estiman el impacto económico de la desnutrición para el año 2011 en dos horizontes de tiempo (retrospectiva y prospectiva). En el trabajo se considera también los costos asociados a la educación y productividad en la economía del Perú. Los datos son extraídos de la Encuesta Nacional de Hogares, el censo de la población y vivienda del año 2007 y la encuesta demográfica de salud familiar. Las recomendaciones que realizan

los autores se centran en los numerosos riesgos a los que están expuestos los niños a causa de la desnutrición y como esta repercute en la productividad en el país.

Rivera y Currais (2014) analizaron la repercusión de la inversión pública en la salud, la productividad de los individuos y el crecimiento económico. El objetivo general era identificar el impacto de la salud en los salarios en España. Los autores encontraron dos problemas metodológicos al realizar la investigación, el primero era la dificultad en el error del estado de salud, pues indicaban que a diferencia de otras formas de capital humano, algunos indicadores de salud no son objetivos, es así que la salud puede considerarse como una variable no observable que se presenta a través de indicadores imperfectos, como los días de enfermedad, los días que un individuo se encontraba limitado a realizar su actividad principal o el padecimiento de una enfermedad, en ese sentido, con la finalidad de evitar heterocedasticidad en el análisis por la medición en las encuestas basada en la percepción personal del individuo acerca de lo que realmente es encontrarse enfermo, utilizaron una variable dicotómica en la que si el individuo se vio limitado a realizar sus actividades principales debido a alguna enfermedad y el número de días en los que el individuo trabajó teniendo problemas de salud en los últimos 15 días. El segundo problema encontrado fue la simultaneidad entre la productividad y salud, el cual podría distorsionar las estimaciones sobre la base de la correlación de los términos de errores. Los resultados obtenidos fueron que la salud tiene un efecto en el crecimiento económico.

Sánchez y Fernández (2016) estudian la pérdida de productividad laboral atribuible al consumo de tabaco, siguiendo las afirmaciones de la Organización Mundial de la Salud, la cual considera que una de las causas de la pérdida de productividad laboral en una economía nacional es por el consumo de tabaco. Los autores tienen como objetivo identificar los fundamentos y características del tabaquismo como un determinante de la pérdida de productividad laboral y para ello realizaron una investigación bibliográfica y un estudio descriptivo sobre el determinante mencionado, los referidos autores encontraron que el tabaquismo puede considerarse un factor de riesgo y clave en la pérdida de productividad laboral, además de impactar negativamente en la economía social e individual.

Domínguez et al. (2022) analizaron el impacto del ausentismo laboral en la productividad y presentaron estrategias de control desde los programas de salud que

podrían tener las empresas con la finalidad de confrontarlo. Los resultados encontrados en base a revisión de literatura son que el ausentismo es un problema que genera altas pérdidas económicas y la población que labora y que cuenta con mayor prevalencia de enfermedades crónicas como diabetes, hipertensión arterial, entre otras, muestran un mayor número de días perdidos. Las estrategias para el control que recomiendan son implementar programas de salud y seguridad en el trabajo, con el objetivo de reducir el impacto de situaciones adversas sobre la productividad e incrementar la calidad de vida de los trabajadores.

Lozano y Riaño-Casallas (2018) investigan cómo la política pública de seguridad y salud en el trabajo construyen una cultura de entornos laborables saludables y productivos. Para lo cual presentan directrices internacionales postuladas para la OMS y la OIT e intentan modelarlos mediante posibles programas de seguridad y salud en el país de Colombia. Concluyen en que el país se encuentra en el momento preciso para evaluar sus políticas públicas y programas nacionales debido a que el país se encuentra siguiendo desde hace 15 años un marco estratégico mundial que le permite considerar una evaluación y posible modificación en la política pública que se aplica.

CAPÍTULO III: METODOLOGÍA

3.1 Datos y variables

La muestra considerada para este estudio comprende 36.994 hogares, correspondiendo 23.346 al área urbana y 13.648 al área rural del Perú; además, en términos individuales se tiene a 22.061 personas encuestadas que cumplen con las condiciones de nuestro estudio. La data corresponde a la Encuesta Nacional de Hogares sobre Condiciones de Vida y Pobreza para el año 2019. Todos los datos fueron extraídos de la página web del INEI.

Las variables utilizadas en este estudio han sido utilizadas en trabajos de investigación previos como Cortez (1999) y Ribero (2000). A continuación, se describen tanto a la variable dependiente, variable independiente y las variables de control que se utilizarán en el presente trabajo de investigación. Cabe resaltar, que dado el rezago que existe respecto a esas dos investigaciones, ahora contamos con más indicadores pertinentes que son considerados dentro de nuestras variables.

La variable dependiente que será utilizada en el modelo econométrico es el logaritmo natural del ingreso mensual, calculado a partir de las horas de trabajo de los individuos y expresado en nuevos soles. Por otro lado, la variable independiente es el indicador de salud y es calculado mediante la respuesta de sí o no con respecto a la pregunta: ¿se ha encontrado enfermo en las últimas 4 semanas? La pregunta como la respectiva respuesta se encuentra en la base de datos con la que se está trabajando. A continuación, se pasará a describir a las variables de control con las que se trabajará:

VARIABLES DE CONTROL:

- Edad: Son los años cumplidos de los individuos (no se considera fraccionamiento de año).
- Cuadrado de la edad: Se utiliza esta variable con el fin de considerar posibles no linealidades en el impacto de la variable edad sobre los salarios.
- Años de educación: Son los años de estudio de los individuos (se toma en consideración los grados aprobados).

- Cuadrado de los años de educación: Se utiliza esta variable con la finalidad de captar la no linealidad de la variable años de educación con relación a los salarios.
- Años de experiencia: Se define a esta variable como el número de los años de experiencia de los individuos en su centro de labores.
- Cuadrado de los años de experiencia: Esta variable es utilizada para captar la no linealidad de la variable años de experiencia con relación a los salarios.
- Total de horas trabajadas: Se define a esta variable como el número de horas de trabajo mensual de los individuos en su centro de labores.
- Sexo: Es una variable dicotómica (dummy), en donde ser hombre representa igual a 1; por otro lado, ser mujer representa el 0.
- Horas de oferta de agua: Esta variable está definida como las horas de abastecimiento de agua de la red pública durante el último mes.
- Tipo de alumbrado: Es una variable dicotómica, en donde el acceso a la electricidad es igual a “electric”; la ausencia del acceso a los servicios de electricidad es igual a “pase”.
- Baño con desagüe: Es una variable dicotómica, en donde el acceso a la red pública de desagüe del individuo es igual a 1; la ausencia del acceso al sistema de desagüe es igual a 0.
- Fachada con tarrajeo: Es una variable dummy, en la cual si la vivienda cuenta con fachada con tarrajeo es “total”, y si no está descrita como “no está”.
- Teléfono celular: Es una variable dicotómica en la que si el individuo cuenta con dicho artefacto es colocado como “teléfono”, caso contrario se describe en la base datos como “pase”.
- Conexión a TV cable o satelital: Es una variable dicotómica en la que si el individuo cuenta con conexión a tv cable se coloca “conexión”, caso contrario se describe como “pase”.
- Fachada pintada: Es una variable dicotómica en la que si la vivienda del individuo cuenta con una fachada pintada se describe como “totalmen”, caso contrario se describirá como “sin pint”.
- Conexión a internet: Es una variable dicotómica en la que si el individuo cuenta con conexión a internet se describirá como “conexión”, caso contrario se le describirá como “pase”.

- Teléfono fijo: Es una variable dicotómica en la que si el individuo cuenta con teléfono fijo en su vivienda se colocará “teléfono”, caso contrario se le describirá como “pase”.
- Teléfono celular: Es una variable dummy en la que si el individuo cuenta con teléfono celular se colocará “celular”, caso contrario se le describirá como “pase”.

3.2 Operacionalización de las variables

Con relación al indicador de la salud, que es una variable dummy, esperamos que el salario sea mayor conforme la persona no esté incapacitada por enfermedad o dolencia crónica.

Asimismo, se espera que, la edad tenga un impacto positivo sobre los salarios, ya que generalmente el salario de una persona se va incrementando a lo largo del tiempo por la experiencia que posee. Además, planteamos que la relación entre el salario y el cuadrado de la edad sería negativa dado que esto representa la no linealidad de la relación entre edad y experiencia, es decir, la segunda derivada de la función de la experiencia será negativa con respecto a la edad y presenta un comportamiento decreciente, por tanto, la relación final con el salario sería negativa.

Con relación a los años de educación, planteamos que, a más años de educación, los salarios se incrementen debido a la mayor capacitación con la que cuenta el individuo; además, esperamos una relación negativa con el cuadrado de la educación, ya que nuevamente esta variable capta la no linealidad que existe entre la educación y los salarios, por tanto, a lo largo del tiempo esta relación se muestra como una curva con primera derivada positiva. pero segunda derivada negativa, por tanto, la relación con el cuadrado de la educación sería negativa.

Con respecto a los años de experiencia, argumentamos que, a más años de experiencia, los salarios se incrementen debido a que los individuos como parte del capital humano posean mayores habilidades, y por ende su desempeño en el ámbito laboral será más desarrollado que los individuos que recién se encuentren iniciando; además, esperamos una relación negativa con el cuadrado de los años de experiencia, ya que esta última capta la no linealidad que existe entre años de experiencia y salarios, por

lo tanto, al igual que la variable cuadrada de la educación que estamos tomando anteriormente, se mostraría una relación con una curva con primera derivada positiva, pero segunda derivada negativa.

Ahora, con respecto a la variable sexo, esperamos una relación positiva tanto para hombres como mujeres; sin embargo, el impacto de los salarios será mayor para las mujeres que los hombres. Esta última afirmación es basada en trabajos anteriores y la cual esperamos confirmar de acuerdo con el contexto que estamos estudiando.

Con respecto a las variables dicotómicas: tipo de alumbrado, acceso a un teléfono celular, acceso a un teléfono fijo, conexión a TV cable y conexión a internet, esperamos una relación positiva, ya que el acceso de estos servicios por parte de los individuos le permitirá contar con un mejor nivel de vida y protección a su salud.

Asimismo, esperamos una relación positiva con respecto a las variables fachada como tarrajeo, fachada pintada y baño con desagüe, debido a que estas variables son indicadores de mejor infraestructura de la vivienda, lo que les permitirá a los individuos mayor protección a su salud, a la vez que un mejor estilo de vida.

Por último, esperamos una relación positiva en cuanto al número total de horas trabajadas y número de horas de oferta de agua, ya que a un mayor de número de horas trabajadas se espera un mayor salario y, por lo tanto, se espera un mayor número de horas del servicio de agua.

Tabla 3.1

Operacionalización de las variables

Dimensión	Notación	Signo Esperado	Fuente
V. DEPENDIENTE			
Ingreso Mensual	(IM)		ENAH0 (2019)
V. INDEPENDIENTE			
Indicador de Salud	(IDS)	(+)	ENAH0 (2019)
V.CONTROL			
Edad	(E)	(+)	ENAH0 (2019)
Cuadrado de la edad	(E2)	(-)	ENAH0 (2019)

Dimensión	Notación	Signo Esperado	Fuente
Número de años de educación	(ED)	(+)	ENAH0 (2019)
Cuadrado de los años de educación	(ED2)	(-)	ENAH0 (2019)
Años de experiencia	(EX)	(+)	ENAH0 (2019)
Cuadrado de los años de experiencia	(EX2)	(-)	ENAH0 (2019)
Sexo	(S)	(+)	ENAH0 (2019)
Horas de oferta de agua	(OA)	(+)	ENAH0 (2019)
Acceso a la electricidad	(AE)	(+)	ENAH0 (2019)
Sistema de desagüe adecuado	(SD)	(+)	ENAH0 (2019)
Total horas trabajadas	(TH)	(+)	ENAH0 (2019)
Fachada con tarrajeo	(FT)	(+)	ENAH0 (2019)
Fachada pintada	(FP)	(+)	ENAH0 (2019)
Tipo de alumbrado	(TA)	(+)	ENAH0 (2019)
Baño con desagüe	(BD)	(+)	ENAH0 (2019)
Conexión a T. V	(CTV)	(+)	ENAH0 (2019)
Conexión a internet	(CI)	(+)	ENAH0 (2019)
Teléfono fijo	(TF)	(+)	ENAH0 (2019)
Teléfono celular	(TC)	(+)	ENAH0 (2019)

Fuente: Elaboración propia.

3.3 Propuesta de estimación empírica

Dado que el objetivo de esta investigación es verificar el impacto de la salud sobre la productividad a nivel desagregado entre diversos criterios como los quintiles de ingreso, dominios geográficos, rangos de edad y niveles educativos; entonces, se realizarán

diversos modelos para cada rubro de cada nivel de desagregación aplicando el método de corrección en dos etapas de Heckman (1979) que corrige el sesgo de selección.

Este modelo, al aplicar una ratio que corrige el sesgo de selección, brinda mayores ventajas sobre un modelo típico de MCO, ya que incorpora los efectos de un vector de variables que permiten un mejor estimador y resultados no sesgados en el modelo econométrico final.

Este método fue aplicado Cortez (1999) cuando estima la ecuación salarial para verificar el impacto de la salud sobre la productividad, el autor emplea el modelo de Heckman en dos etapas debido a las bondades que presenta para corregir el sesgo de selección. En conclusión, el método propuesto por Heckman es preciso para esta investigación debido a que permite corregir el sesgo de selección que se produce cuando la muestra a disposición no es “aleatoria” y, por ende, puede no estar representando adecuadamente la población que se desea estudiar.

En primera instancia se estima un modelo probit para calcular la probabilidad de que el individuo forma parte del mercado laboral, en esta ecuación se incluyen las variables explicativas como el salario, la salud y demás variables que se detallan en la matriz de consistencia. Además, se calcula el denominado “inverso de la ratio de Mills” (λ_i) que se incorporará en la segunda etapa. En esta primera etapa se parte de incorporar la probabilidad de que el sujeto muestreado tenga empleo, lo cual se muestra en la siguiente forma matemática:

$$Prob(D = 1|Z) = \phi(Z\gamma)$$

Respecto a lo anterior, si D es igual a 1 el muestreado si se encuentra empleado y si es 0 entonces no se encuentra empleado, Z es el vector de variables explicativas que se están considerando, γ es un vector de parámetros desconocidos y ϕ denota la función de distribución acumulativa. Ahora bien, luego de realizar esta estimación mediante un modelo probabilístico, que puede ser logit o probit, se procede a la segunda etapa.

En la segunda etapa se incorpora el factor hallado en la primera etapa como un regresor del modelo MCO calculado en esta segunda etapa, de esta forma la significatividad de esta variable incluida indica la magnitud del sesgo en que se incurriría si no se hubiera incluido la variable. Este factor que se incorpora tiene la siguiente forma:

$$\omega^* = X\beta + u$$

Con base a ello, en la ecuación de salarios que se muestra en la parte superior, ω^* denota la variable salario que va a estar a su vez vinculado a la siguiente forma matemática de la probabilidad que el muestreado se encuentre empleado:

$$E[\omega|X, D = 1] = X\beta + E[u|X, D = 1]$$

Lo anterior describe la condicionalidad de la variable salario que, como mencionamos, está vinculado a la esperanza condicional de que el sujeto muestreado trabaja. Por último, la ecuación final que incorpora la variable “inverso de la ratio de Mills” (λ), entonces, tendríamos finalmente la ecuación:

$$E[\omega|X, D = 1] = X\beta + \rho\sigma_u\lambda(Z_\gamma)$$

La forma de la ecuación superior es la que corresponde a la regresión de la segunda etapa y donde al incorporar el factor λ obtenemos la corrección del sesgo de selección. En ese sentido, dado que queda claro que nuestra metodología será la de Heckman en 2 etapas, la estructura final de nuestro modelo econométrico viene dado por:

$$IM_{it} = \alpha + \beta_0 \cdot IDS_{it} + \beta_1 \cdot E_{it} + \beta_2 \cdot E2_{it} + \beta_3 \cdot ED_{it} + \beta_4 \cdot ED2_{it} + \beta_5 \cdot EX_{it} + \beta_6 \cdot EX2_{it} + \beta_7 \cdot S_{it} + \beta_8 \cdot OA_{it} + \beta_9 \cdot TF_{it} + \beta_{10} \cdot BD_{it} + \beta_{11} \cdot CI_{it} + \beta_{12} \cdot THT_{it} + \beta_{13} \cdot imr_{it} + \varepsilon_{it}$$

Donde:

i : Corresponde a las 121.623 personas encuestadas que son consideradas en el estudio.

t : Año 2019 (frecuencia anual).

IM : Ingreso Mensual.

IDS : Indicador de Salud.

E : Edad.

$E2$: Cuadrado de la edad.

ED : Número de años de educación.

ED2: Cuadrado de los años de educación.

EX: Experiencia.

EX2: Cuadrado de los años de experiencia.

S: Sexo.

OA: Horas de Oferta de Agua.

TF: Acceso a Telefonía Fija.

BD: Sistema de desagüe adecuado.

CI: Conexión a Internet.

THT: Total de Horas Trabajadas.

imr: Inverso del Ratio Mills.

ε : Error

α : Termino independiente.

3.4 Estadísticas Descriptivas

A continuación, se presentan las estadísticas obtenidas de la ENAHO (2019), las cuales consideran las medidas de distribución porcentual entre dominios geográficos, rangos de edad, quintiles de ingreso y niveles educativos, así como el indicador de salud y promedio de ingreso asociados a estos cuatro niveles de desagregación.

Para el análisis inicial hemos considerado las estadísticas generales que comparten todos los individuos, las cuales hemos dividido en 2 cuadros descriptivos. El primer cuadro considera características individuales de cada infórmate que conforma nuestra muestra de 22.061 encuestados. Luego se presentan las características asociadas a las condiciones de la vivienda de los informantes; esto es, el acceso a servicios públicos, características externas asociadas a la fachada y otros servicios adicionales.

Finalmente, se presenta el análisis más concreto sobre las 4 desagregaciones que estamos considerando en la presente investigación, además, para brindar mayor relevancia a nuestros resultados hemos incorporado la prueba de diferencia de medias.

Tabla 3.2

Características de la muestra de encuestados desagregado por el indicador de salud y género

Variable	Sí estuvo enfermo em las últimas 4 semanas				No estuvo enfermo em las últimas 4 semanas				Total	Test de Diferencia de Medias	
	Hombre		Mujer		Hombre		Mujer			Hombre	Mujer
	Media	Desv. Est.	Media	Desv. Est.	Media	Desv. Est.	Media	Desv. Est.	Media	Hombre	Mujer
Años de educación	11.25	3.95	11.74	4.19	10.96	4.07	11.70	4.19	11.31	3.46***	0.34
Edad	38.13	13.45	37.33	12.62	38.44	13.72	38.09	12.66	38.17	1.13	2.59***
Ingresos	1149.55	1151.16	1007.17	1009.69	1166.32	1252.76	1049.83	1107.50	1110.68	0.66	1.68**
Horas Trabajadas	43.16	17.64	37.37	17.46	43.32	17.65	38.61	17.41	41.20	0.44	3.034***
Experiencia	5.33	8.32	5.12	8.53	5.85	9.12	5.26	8.71	5.52	2.82***	0.69
Número de observaciones	3078		2492		9801		6690		22061		

La muestra incluye trabajadores entre 18 y 70 años.

(*) Significancia al 10% de confianza; (**) Significancia al 5% de confianza; (***) significancia al 1% de confianza.

Fuente: INEI (2019). *Encuesta Nacional de Hogares (ENAHO)*. Recuperado el 24 de septiembre de 2020, de Instituto Nacional de Estadística e Informática:

<http://iinei.inei.gob.pe/microdatos/>. Elaboración propia.

En la tabla 3.2 se muestran las estadísticas descriptivas sobre las características individuales de los 22 061 encuestados considerados en la presente investigación. En ese sentido, se puede observar que los individuos en promedio tienen 11.31 años de educación y que tanto hombres (10.96) como mujeres (11.70) que no presentaron enfermedad crónica en las últimas 4 semanas tienen en promedio menos años de educación que los hombres (11.25) y mujeres (11.74) que sí estuvieron enfermos en las últimas 4 semanas. Cabe precisar que el resultado anterior es significativo al 1% para el caso de los hombres y para el caso de las mujeres no es significativo. Ahora bien, lo anterior no es algo esperado; sin embargo, una posible explicación puede ser que los hombres que estudiaron más tengan una mayor cantidad de trabajo y producto de ello estén más propensos a sufrir enfermedades o dolencias

En cuanto a la edad, tenemos que los individuos en promedio tienen 38.17 años y que en promedio tanto hombres (38.13) como mujeres (37.33) que han presentado alguna enfermedad en las últimas 4 semanas son menores que los hombres (38.44) y mujeres (38.09) que no la han presentado. Cabe precisar que el resultado anterior es significativo al 1% para el caso de las mujeres y para el caso de los hombres el resultado no es significativo. Lo anterior tampoco es algo que se esperaba; sin embargo, una posible explicación puede ser que las más jóvenes cuentan con menos recursos para acceder a mayores cuidados de salud.

Respecto de los ingresos, vemos que en promedio los individuos perciben 1 110.68 soles; además, comprobamos que tanto los hombres (S/1 166.32) como mujeres (S/1 049.83) que no han presentado enfermedad en las últimas 4 semanas presentan mayores ingresos que aquellos hombres (S/1 149.55) y mujeres (S/1 007.17) que sí han presentado enfermedad. Cabe precisar que el resultado anterior es significativo al 5% para el caso de las mujeres y para el caso los hombres el resultado no es significativo. Esto último es de suma importancia para nuestro estudio, pues se comprueba parcialmente que aquellas personas que están enfermas tienen menores ingresos que aquellas personas que no han estado enfermas.

Ahora bien, en cuanto a las horas trabajadas, vemos que en promedio se trabaja 41.20 horas a la semana; además, notamos que los hombres (43.32) y mujeres (38.61) que no estuvieron enfermos en las últimas 4 semanas trabajan en promedio más horas que los hombres (43.16) y mujeres (37.37) que sí estuvieron enfermos. Cabe precisar que el resultado anterior es significativo al 1% para el caso de las mujeres y para el caso los

hombres el resultado no es significativo. Lo anterior se puede deber a que las personas que trabajan más horas corresponden con aquellas personas que cuentan con un trabajo de horario completo que goza de los beneficios de ley como el acceso al seguro que permite realizar mayores chequeos de salud y prevenir enfermedades.

Por último, en cuanto a la experiencia, en promedio los individuos tienen 5.52 años de experiencia y se puede notar que tanto hombres (5.85) como mujeres (5.26) que no estuvieron enfermos cuentan con mayor experiencia que aquellos hombres (5.33) y mujeres (5.12) que sí estuvieron enfermos en las últimas 4 semanas. Cabe precisar que el resultado anterior es significativo al 1% para el caso de los hombres y para el caso las mujeres el resultado no es significativo. Esto último se esperaba, ya que las personas con mayor experiencia pueden acceder a mejores trabajos que brinden mayores beneficios como mejores servicios de salud.

En la tabla 3.3 se muestran las estadísticas descriptivas sobre las características asociadas al hogar de los 22 061 encuestados considerados en la investigación. Respecto a estas características, podemos notar que aquellas personas que no estuvieron enfermas en las últimas 4 semanas tenían en promedio mayor acceso a los servicios de: agua potable (13.7212), televisión por cable o satelital (46.14%) y telefonía fija (20.62%) en comparación a aquellas personas que sí presentaron enfermedad y tenían menor acceso a los servicios de: agua potable (13.7176), televisión por cable o satelital (45.37%) y telefonía fija (19.07%). Ello nos muestra un impacto positivo de estos servicios para la mayor prevención de las enfermedades entre los ciudadanos, por una parte, el agua potable contribuye a la higiene, la televisión por cable o satelital contribuye a difundir las campañas de prevención sobre alguna enfermedad y la telefonía fija permite mayor comunicación con los centros de salud sobre alguna consulta médica.

También podemos observar que aquellas personas que sí estuvieron enfermas en las últimas 4 semanas tenían en promedio mayor acceso a los servicios de: alumbrado eléctrico (97.72%), internet (47.43%) y desagüe (77.67%) en comparación a aquellas personas que no presentaron enfermedad y tenían menor acceso a los servicios de: alumbrado eléctrico (97.40%), internet (46.46%) y desagüe (76.47%). Lo anterior no es algo que se esperaba debido a que el mayor acceso a desagüe permite mayor higiene del hogar, el internet permite mayor rapidez en la comunicación de campañas y el alumbrado eléctrico permite mayor acceso a los equipos electrónicos como laptop y televisión.

Tabla 3.3

Características del hogar sobre la muestra de encuestados desagregado por el indicador de salud

Variable	Sí estuvo enfermo en las últimas 4 semanas		No estuvo enfermo en las últimas 4 semanas		Total
	Media	% del total	Media	% del total	Media / % del total
Horas de Oferta de Agua	13.7176		13.7212		13.7203
Alumbrado Eléctrico	5443	97.72%	16062	97.40%	97.48%
Acceso Celular	5474	98.28%	16103	97.65%	97.81%
Acceso a TV Cable/Satelital	2527	45.37%	7609	46.14%	45.95%
Acceso Internet	2642	47.43%	7662	46.46%	46.71%
Teléfono Fijo	1062	19.07%	3401	20.62%	20.23%
Fachada con Tarrajeo	3020	54.22%	9102	55.19%	54.95%
Fachada Pintada	2927	52.55%	8931	54.16%	53.75%
Baño con Desagüe	4326	77.67%	12610	76.47%	76.77%
Número de Observaciones	5570	100.00%	16491	100.00%	100.00%

La muestra incluye trabajadores entre 18 y 70 años.

Fuente: INEI (2019). *Encuesta Nacional de Hogares (ENAH0)*. Recuperado el 24 de septiembre de 2020, de Instituto Nacional de Estadística e Informática: <http://inei.inei.gob.pe/microdatos/>. Elaboración propia.

Por último, la tabla 3.3 muestra que, respecto a las características exteriores de la vivienda, las personas que no presentaron enfermedad en las últimas 4 semanas tienen en promedio sus viviendas más tarrajeadas (55.19%) y pintadas (54.16%) que las personas que sí presentaron enfermedad y tienen sus casas menos tarrajeadas (54.22%) y pintadas (52.55%). Esto último, nos muestra que las características exteriores de la vivienda pueden ser considerados como variables relevantes para inferir sobre el estado de salud del hogar.

La tabla 3.4 nos muestra el ingreso promedio por sexo, dominio geográfico y correspondencia con el indicador de salud. En ese sentido, la primera observación en la tabla es que para los 6 dominios geográficos: Costa centro, Costa norte, Lima Metropolitana, Selva, Sierra centro y Sierra sur se cumple la condición de que las personas que no estuvieron enfermas en las últimas 4 semanas tienen mayores ingresos promedio que aquellas personas que sí tuvieron alguna enfermedad. Por otro lado, se observa que para los dominios Costa sur y Sierra norte, el efecto de la salud es contrario al esperado.

Por otro parte, se observa también que en Lima Metropolitana se refleja el mayor efecto del indicador de salud sobre el ingreso promedio, esto no se esperaba, ya que en Lima Metropolitana generalmente se tiene mayor acceso a los servicios de salud y se espera que las personas cuenten con un mejor estado de salud y mejores rendimientos salariales; en la contraparte tenemos a la Selva donde se muestra el menor impacto de la salud sobre los ingresos promedio. Con relación a estos resultados que se observan en Lima Metropolitana y la Selva, podemos notar que el factor salud puede estar afectando en mayor proporción a los dominios donde hay mayor concentración de población y, por ende, a pesar de que Lima Metropolitana cuente con mayores centros de salud, las personas tienen dificultades para acceder a los servicios de salud debido a la gran cantidad de personas que demandan este servicio, es por ello que los ingresos se vean más afectados; por el contrario, en la Selva se pueden tener menos centros de salud pero mayor accesibilidad a los centros de salud debido a la menor concentración de población.

Tabla 3.4

Ingreso promedio desagregado por dominios geográficos, indicador de salud y sexo

Dominios Geográficos	Sí estuvo enfermo em las últimas 4 semanas	No estuvo enfermo em las últimas 4 semanas	Total General	Test de Diferencia de Medias
Costa Centro	877.633	909.817	899.918	0.905
<i>Mujer</i>	834.586	861.360	852.492	0.508
<i>Hombre</i>	914.429	943.955	935.397	0.613
Costa Norte	859.463	915.276	906.325	1.358*
<i>Mujer</i>	821.049	912.970	896.517	1.526*
<i>Hombre</i>	891.163	916.788	913.004	0.457
Costa Sur	1280.574	1201.956	1218.616	1.095
<i>Mujer</i>	1072.457	949.382	975.479	1.539*
<i>Hombre</i>	1458.021	1417.018	1425.703	0.367
Lima Metropolitana	1333.839	1534.280	1484.839	3.316***
<i>Mujer</i>	1271.986	1460.438	1411.370	2.268***
<i>Hombre</i>	1388.496	1591.439	1543.607	2.346***
Selva	1017.842	1021.815	1020.911	0.100
<i>Mujer</i>	978.883	1052.231	1034.379	1.220
<i>Hombre</i>	1046.923	1002.267	1011.952	0.850
Sierra Central	1043.792	1050.996	1048.692	0.173
<i>Mujer</i>	897.458	904.162	901.729	0.113
<i>Hombre</i>	1173.027	1143.690	1152.178	0.512
Sierra Norte	1178.416	1012.848	1059.089	2.147**
<i>Mujer</i>	1202.208	1084.590	1117.866	0.959

<i>Hombre</i>	1163.940	970.432	1024.058	1.951
Sierra Sur	1219.224	1312.025	1281.352	1.663**
<i>Mujer</i>	1111.520	1023.512	1054.579	1.309*
<i>Hombre</i>	1296.932	1488.289	1427.818	2.38***

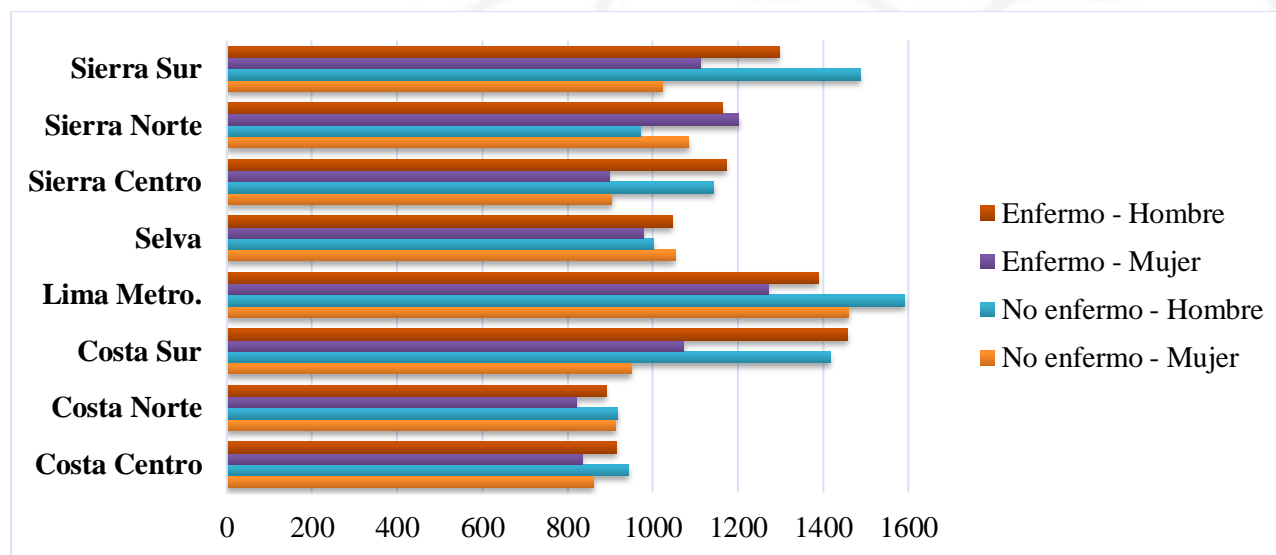
La muestra incluye trabajadores entre 18 y 70 años.

(*) Significancia al 10% de confianza; (**) Significancia al 5% de confianza; (***) significancia al 1% de confianza.

Fuente: INEI (2019). *Encuesta Nacional de Hogares (ENAHO)*. Recuperado el 24 de septiembre de 2020, de Instituto Nacional de Estadística e Informática: <http://inei.inei.gob.pe/microdatos/>. Elaboración propia.

Figura 3.1

Ingreso promedio desagregado por dominios geográficos, indicador de salud y sexo



Fuente: INEI (2019). *Encuesta Nacional de Hogares (ENAHO)*. Recuperado el 24 de septiembre de 2020, de Instituto Nacional de Estadística e Informática: <http://inei.inei.gob.pe/microdatos/>. Elaboración propia.

Finalmente, un hallazgo importante es la relevancia de haber desagregado la región costa y sierra en sus diferentes zonas: norte, centro y sur. Ello nos muestra que la zona sur tanto para la Costa como para la Sierra son las que poseen los mayores ingresos promedio y el mayor efecto de la salud sobre los ingresos. Para la Sierra sur sí se cumple que las personas no enfermas ganen más que las que sí estuvieron enfermas; sin embargo, podemos notar que para la zona sur de la Costa no se cumple que el indicador de salud impacte de forma positiva sobre los ingresos promedio. Mientras que el menor impacto de indicador de salud sobre los ingresos promedio se da en la zona norte de la región Costa y en la zona centro de la región Sierra.

Por otra parte, la tabla 3.5 nos muestra la división entre enfermos y no enfermos que existe para cada dominio geográfico subdividido por el género del entrevistado, ello muestra que la mayor parte de los entrevistados proviene de la Costa Norte y la menor parte de entrevistados proviene de la Sierra Norte, además, en todos los dominios geográficos se cumple que la proporción de hombres encuestados es mayor a la proporción de mujeres encuestadas y que los encuestados no enfermos son más que los encuestados enfermos.

Tabla 3.5

Número de individuos enfermos y no enfermos desagregado por dominios geográficos

Dominios Geográficos	Número de individuos enfermos	Número de individuos no enfermos	Total de individuos
Costa Centro	907 (30.76%)	2042 (69.24%)	2949
<i>Mujer</i>	418 (33.12%)	844 (66.88%)	1262
<i>Hombre</i>	489 (29.00%)	1198 (71.00%)	1687
Costa Norte	637 (16.00%)	3335 (84.00%)	3972
<i>Mujer</i>	288 (17.90%)	1321 (82.10%)	1609
<i>Hombre</i>	349 (14.77%)	2014 (85.23%)	2363
Costa Sur	352 (21.20%)	1309 (78.80%)	1661
<i>Mujer</i>	162 (21.20%)	602 (78.80%)	764
<i>Hombre</i>	190 (21.20%)	707 (78.80%)	897
Lima Metropolitana	923 (24.67%)	2819 (75.33%)	3742
<i>Mujer</i>	433 (26.00%)	1230 (74.00%)	1663
<i>Hombre</i>	490 (23.57%)	1589 (76.43%)	2079
Selva	861 (22.75%)	2924 (77.25%)	3785

Dominios Geográficos	Número de individuos enfermos	Número de individuos no enfermos	Total de individuos
<i>Mujer</i>	368 (24.34%)	1144 (75.66%)	1512
<i>Hombre</i>	493 (21.70%)	1780 (78.30%)	2273
Sierra Central	838 (32.00%)	1783 (68.00%)	2621
<i>Mujer</i>	393 (36.30%)	690 (63.70%)	1083
<i>Hombre</i>	445 (28.90%)	1093 (71.10%)	1538
Sierra Norte	267 (27.90%)	689 (72.10%)	956
<i>Mujer</i>	101 (28.30%)	256 (71.70%)	357
<i>Hombre</i>	166 (27.70%)	433 (72.30%)	599
Sierra Sur	785 (33.00%)	1590 (67.00%)	2375
<i>Mujer</i>	329 (35.30%)	603 (64.70%)	932
<i>Hombre</i>	456 (31.60%)	987 (68.40%)	1443

La muestra incluye trabajadores entre 18 y 70 años.

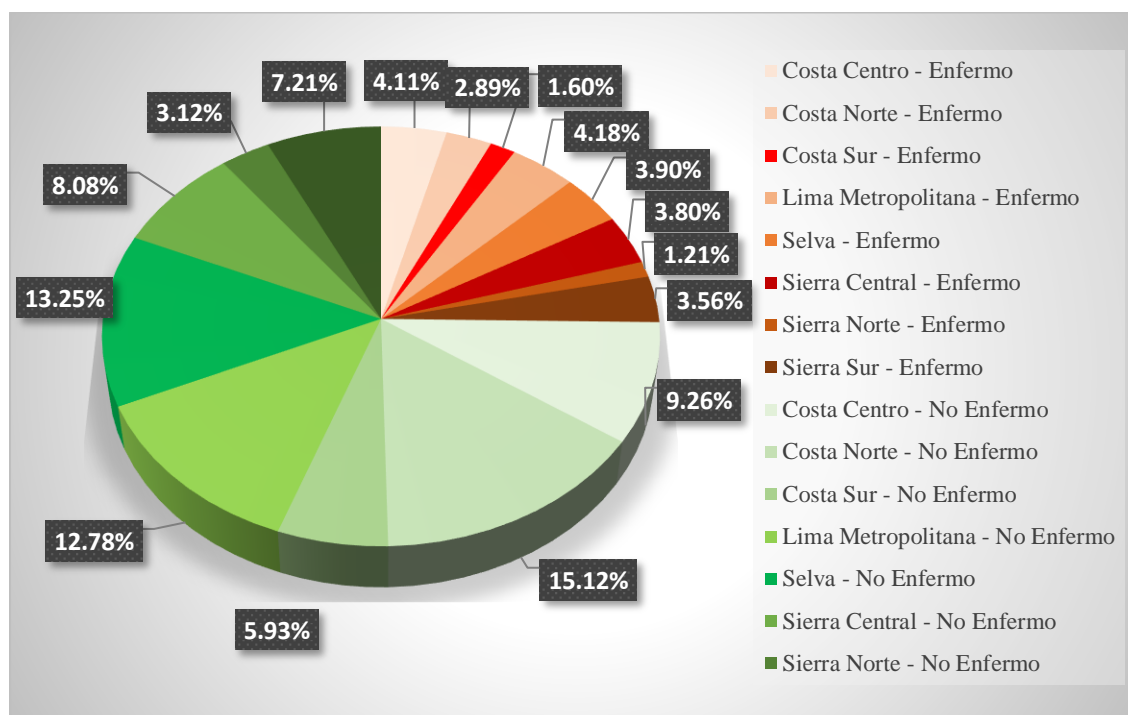
Fuente: INEI (2019). *Encuesta Nacional de Hogares (ENAHO)*. Recuperado el 24 de septiembre de 2020, de Instituto Nacional de Estadística e Informática: <http://inei.inei.gob.pe/microdatos/>. Elaboración propia.

En cuanto a los porcentajes relativos a cada dominio también se detalla en el cuadro 3.5 y notamos que la mayor cantidad de enfermos, en función del total de cada dominio, se halla en la Sierra Sur con 33%; por el otro lado, la menor cantidad se halla en la costa norte con 16%. Ahora bien, lo anterior nos puede indicar que los resultados econométricos sean más consistentes en dominios como la Sierra Sur, donde se presenta mayor muestra de individuos enfermos.

La figura 3.2 nos muestra que la proporción de individuos enfermos representan el 25.25% del total, es decir, prácticamente la cuarta parte de los individuos encuestados tienen la condición de haber estado enfermo en las últimas 4 semanas. Además, respecto al total, la cantidad de encuestados enfermos con menor participación es del dominio Sierra Norte con 1.21% y la de mayor participación es del dominio de Lima Metropolitana con 4.18%. En cuanto a la cantidad de encuestados no enfermos, el dominio con mayor participación es la de Costa Norte con 15.12% y el dominio con menor participación es Sierra Central con 3.12%.

Figura 3.2

Número de individuos enfermos y no enfermos desagregado por dominios geográficos



Fuente: INEI (2019). *Encuesta Nacional de Hogares (ENAHO)*. Recuperado el 24 de septiembre de 2020, de Instituto Nacional de Estadística e Informática: <http://inei.inei.gob.pe/microdatos/>. Elaboración propia.

La tabla 3.6 nos muestra el ingreso promedio por sexo, niveles educativos y correspondencia con el indicador de salud. En ese sentido, la primera observación en la tabla es que para los 3 niveles educativos: Primaria, Secundaria y Superior, se cumple que para las personas que estuvieron alguna enfermedad crónica en las últimas 4 semanas tienen un ingreso promedio menor a las personas que no tuvieron enfermedad alguna. Además, se nota también que el ingreso promedio de las personas con nivel primaria (453.31) son menores que las del nivel secundaria (741.98) y las del nivel superior (1637.38), y el ingreso del nivel secundaria es inferior al del nivel superior. Lo anterior nos muestra que el nivel educativo es un factor significativo para el ingreso promedio y que el impacto del indicador de la salud es el esperado en los 3 niveles educativos considerados.

Tabla 3.6

Ingreso promedio desagregado por niveles educativos, indicador de salud y sexo

Niveles Educativos	Enfermo en las últimas 4 semanas	No enfermo en las últimas 4 semanas	Total General	Test de Diferencia de Medias
Primaria	443.207	456.561	453.313	0.633
<i>Mujer</i>	347.201	362.069	357.895	0.702
<i>Hombre</i>	530.115	517.249	520.039	0.388
Secundaria	726.058	746.960	741.980	1.132
<i>Mujer</i>	546.453	559.610	556.271	0.561
<i>Hombre</i>	828.613	841.075	838.205	0.500
Superior	1547.112	1670.391	1637.376	3.969***
<i>Mujer</i>	1440.547	1545.454	1516.098	2.701***
<i>Hombre</i>	1655.746	1783.510	1750.731	2.668***

La muestra incluye trabajadores entre 18 y 70 años.

(*) Significancia al 10% de confianza; (**) Significancia al 5% de confianza; (***) significancia al 1% de confianza.

Fuente: INEI (2019). *Encuesta Nacional de Hogares (ENAHOG)*. Recuperado el 24 de septiembre de 2020, de Instituto Nacional de Estadística e Informática: <http://inei.inei.gob.pe/microdatos/>. Elaboración propia.

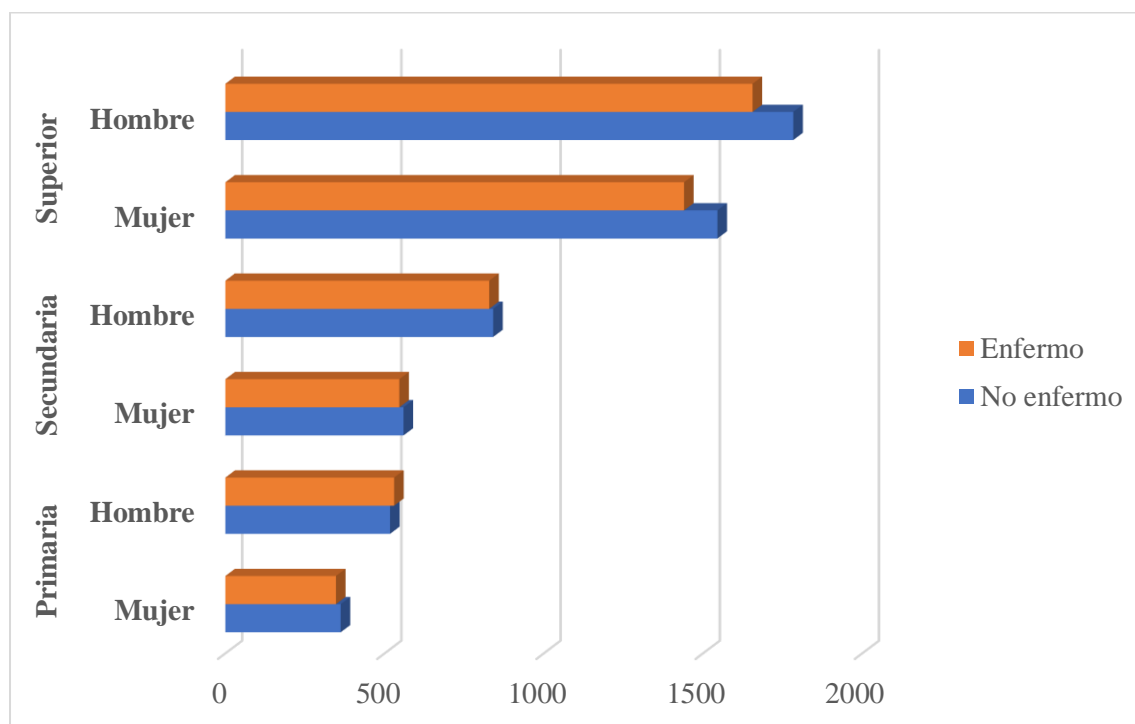
Ahora bien, la figura 3.3 nos muestra que en el nivel superior tanto las mujeres que presentaron alguna enfermedad en las últimas 4 semanas como las que no presentaron enfermedad ganan menos en promedio que los hombres que presentaron enfermedad y los que no la presentaron. Además, vemos que el impacto del indicador de salud es el esperado sobre los ingresos promedio.

En cuanto al nivel secundaria, notamos que al igual que en el nivel superior, tanto las mujeres que presentaron alguna enfermedad en las últimas 4 semanas como las que no presentaron enfermedad ganan menos en promedio que los hombres que presentaron enfermedad y los que no la presentaron. Además, vemos que el impacto del indicador de salud es el esperado sobre los ingresos promedio.

Por último, respecto del nivel primaria, vemos que al igual que el nivel superior y nivel secundaria, tanto las mujeres que presentaron alguna enfermedad en las últimas 4 semanas como las que no presentaron enfermedad ganan menos en promedio que los hombres que presentaron enfermedad y los que no la presentaron. Finalmente, vemos que el impacto del indicador de salud es el esperado sobre los ingresos promedio para el caso de las mujeres; sin embargo, para el caso de los hombres no se cumple tal condición.

Figura 3.3

Ingreso promedio desagregado por niveles educativos, indicador de salud y sexo



Fuente: INEI (2019). *Encuesta Nacional de Hogares (ENAHO)*. Recuperado el 24 de septiembre de 2020, de Instituto Nacional de Estadística e Informática: <http://iinei.inei.gob.pe/microdatos/>. Elaboración propia.

La tabla 3.7 muestra la proporción de individuos enfermos y no enfermos respecto a cada nivel educativo. En primer lugar, notamos que en el nivel educativo superior se presenta la mayor proporción de encuestados, además, en todos los dominios, la cantidad de encuestados hombres es mayor a la de las mujeres. En cuanto al porcentaje relativo de enfermos con respecto a cada nivel educativo, tenemos que la mayor proporción está en el nivel educativo superior con 26.78% y la menor está en el nivel secundaria con 23.83%; además, si bien es cierto que la diferencia entre estas proporciones no es tan grande, la prueba de diferencia de medias calculada en la tabla 3.6 nos brinda grandes indicios a un nivel de significancia del 1% de que el nivel superior va a obtener los resultados econométricos más consistentes en comparación al nivel primaria o secundaria donde la prueba de diferencia de medias resulta no significativa.

Tabla 3.7

Número de individuos enfermos y no enfermos desagregado por niveles educativos

Niveles Educativos	Número de individuos enfermos	Número de individuos no enfermos	Total de individuos
Primaria	764 (24.30%)	2378 (75.70%)	3142
<i>Mujer</i>	363 (28.10%)	930 (71.90%)	1293
<i>Hombre</i>	401 (21.70%)	1448 (78.30%)	1849
Secundaria	2102 (23.83%)	6720 (76.17%)	8822
<i>Mujer</i>	764 (25.37%)	2247 (74.63%)	3011
<i>Hombre</i>	1338 (23.00%)	4473 (77.00%)	5811
Superior	2704 (26.78%)	7393 (73.22%)	10097
<i>Mujer</i>	1365 (28.00%)	3513 (72.00%)	4878
<i>Hombre</i>	1339 (25.66%)	3880 (74.34%)	5219

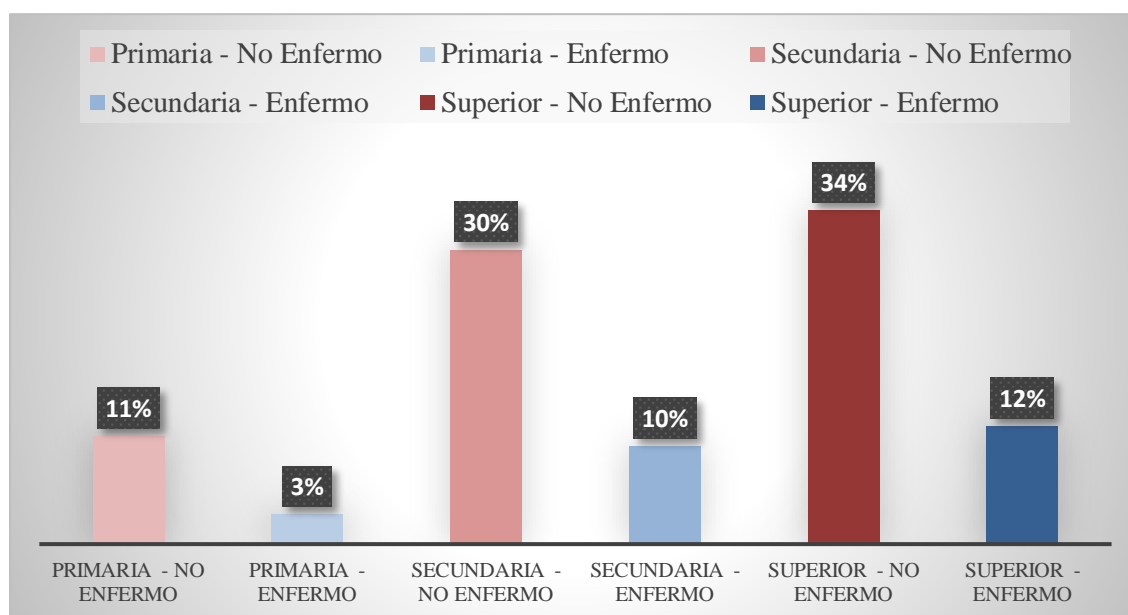
La muestra incluye trabajadores entre 18 y 70 años.

Fuente: INEI (2019). *Encuesta Nacional de Hogares (ENAHO)*. Recuperado el 24 de septiembre de 2020, de Instituto Nacional de Estadística e Informática: <http://inei.inei.gob.pe/microdatos/>. Elaboración propia.

La figura 3.4 nos muestra que la proporción de individuos enfermos representan el 25% del total, es decir, la cuarta parte de los individuos encuestados tienen la condición de haber estado enfermo en las últimas 4 semanas. Además, respecto al total, la cantidad de encuestados enfermos con menor participación es del nivel educativo primaria con 3% y la de mayor participación es del nivel educativo superior con 12%. En cuanto a la cantidad de encuestados no enfermos, el nivel educativo con mayor participación es el superior con 34% y el nivel educativo con menor participación es el nivel primario con 11%.

Figura 3.4

Número de individuos enfermos y no enfermos desagregado por niveles educativos



Fuente: INEI (2019). *Encuesta Nacional de Hogares (ENAHOG)*. Recuperado el 24 de septiembre de 2020, de Instituto Nacional de Estadística e Informática: <http://inei.inei.gob.pe/microdatos/>. Elaboración propia.

La tabla 3.8 nos muestra el ingreso promedio por sexo, niveles educativos y correspondencia con el indicador de salud. La primera observación en la tabla es que, para todos los quintiles, excepto para el tercer quintil, se cumple que las personas que tuvieron alguna enfermedad crónica en las últimas 4 semanas tienen un ingreso promedio menor a las personas que no tuvieron enfermedad alguna. Además, se nota también que tanto los hombres como las mujeres que estuvieron enfermos en las últimas 4 semanas tienen un ingreso promedio menor que los hombres y mujeres que no tuvieron alguna enfermedad.

La única excepción en lo anterior mencionado es en los hombres pertenecientes al tercer quintil, en donde los hombres que estuvieron enfermos (827,335) tienen un ingreso promedio mayor que los hombres que no estuvieron enfermos (821,798). Finalmente, destacamos en cuanto a la prueba de diferencia de medias, calculado también en la tabla 3.8, que el quintil más representativo es el primero con 1% de significancia, seguido del quinto, cuarto, y segundo quintil respectivamente con 5% de significancia, mientras que el único quintil que se muestra como no significativo es el tercero.

Tabla 3.8

Ingreso promedio desagregado por quintiles de ingreso, indicador de salud y sexo

Quintiles de Ingreso	Estuvo enfermo en las últimas 4 semanas	No estuvo enfermo en las últimas 4 semanas	Total General	Test de Diferencia de Medias
Primer Quintil	116.345	122.664	121.086	3.1418***
<i>Mujer</i>	107.580	112.393	111.054	1.675**
<i>Hombre</i>	125.943	131.125	129.960	1.8761**
Segundo Quintil	344.253	348.978	347.798	1.5139*
<i>Mujer</i>	355.429	365.021	362.493	1.8141**
<i>Hombre</i>	336.993	339.698	339.044	0.7114
Tercer Quintil	825.544	824.008	824.417	0.3019
<i>Mujer</i>	823.900	826.451	825.725	0.3590
<i>Hombre</i>	827.335	821.798	823.174	0.7578
Cuarto Quintil	1355.756	1371.094	1367.343	2.0003**
<i>Mujer</i>	1330.203	1343.924	1340.356	1.0888
<i>Hombre</i>	1371.665	1385.946	1382.578	1.4859*
Quinto Quintil	2792.942	2925.861	2892.338	2.7841**
<i>Mujer</i>	2716.485	2810.683	2785.739	1.4118*
<i>Hombre</i>	2846.009	2997.550	2960.552	2.3049**

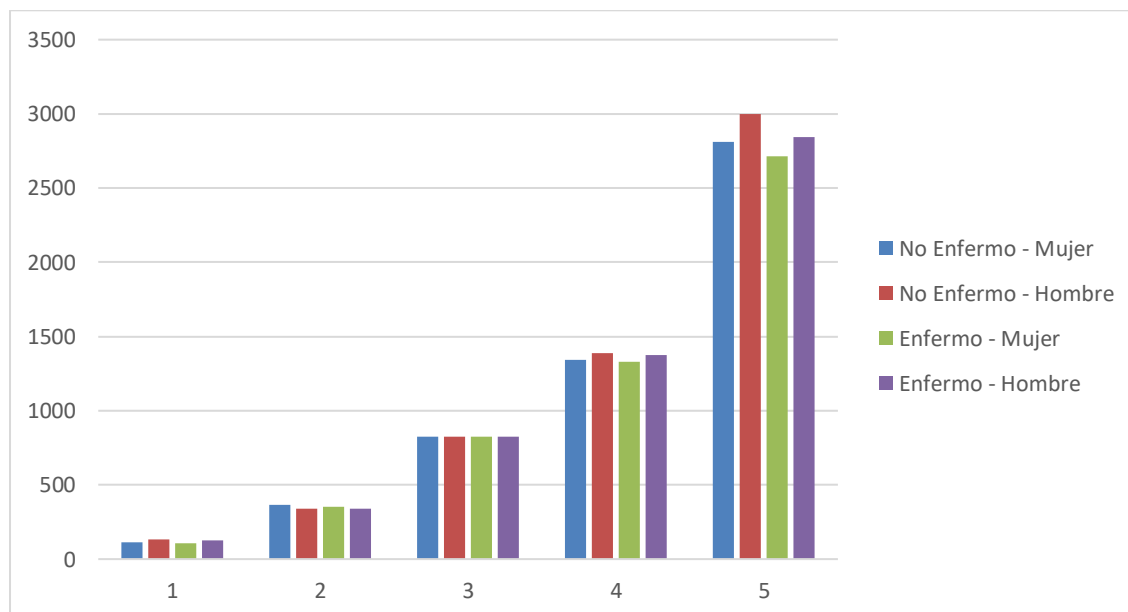
*La muestra incluye trabajadores entre 18 y 70 años.

Fuente: INEI (2019). *Encuesta Nacional de Hogares (ENAH)*. Recuperado el 24 de septiembre de 2020, de Instituto Nacional de Estadística e Informática: <http://inei.inei.gob.pe/microdatos/>. Elaboración propia.

Ahora bien, la figura 3.5 nos muestra que el impacto del indicador de salud es el esperado para los hombres pertenecientes a todos los quintiles, excepto los que pertenecen al tercer quintil. Por otro lado, el impacto de la salud es el esperado para las mujeres pertenecientes a todos los quintiles de ingreso.

Figura 3.5

Ingreso promedio desagregado por quintiles de ingreso, indicador de salud y sexo



Fuente: INEI (2019). *Encuesta Nacional de Hogares (ENAHO)*. Recuperado el 24 de septiembre de 2020, de Instituto Nacional de Estadística e Informática: <http://inei.inei.gob.pe/microdatos/>. Elaboración propia.

La Tabla 3.9 muestra la división entre enfermos y no enfermos, subdividido por el género del entrevistado, que existe para cada quintil de ingreso. Lo primero que se puede verificar es que cada quintil de ingreso contiene la misma cantidad de individuos, que en este caso es de 4412, por lo que la cantidad de entrevistados está repartida de una manera equitativa en esta desagregación. También se puede visualizar que en los entrevistados no enfermos, la proporción de hombres encuestados es mayor a la proporción de mujeres encuestadas; mientras tanto, en los entrevistados enfermos se cumple esa misma razón de que la proporción de hombres encuestados es mayor a la proporción de mujeres encuestadas, excepto en el primer y tercer quintil, en donde las mujeres encuestadas son mayores en proporción que los hombres encuestados.

Finalmente, se muestra que los individuos no enfermos encuestados son más que los encuestados enfermos.

Tabla 3.9

Número de individuos enfermos y no enfermos desagregado por quintiles de ingreso

Quintiles de Ingreso	Número de individuos enfermos	Número de individuos no enfermos	Total de individuos
Primer Quintil	1102 (24.98%)	3310 (75.02%)	4412
<i>Mujer</i>	576 (27.81%)	1495 (72.19%)	2071
<i>Hombre</i>	526 (22.47%)	1815 (77.53%)	2341
Segundo Quintil	1102 (24.98%)	3310 (75.02%)	4412
<i>Mujer</i>	434 (26.35%)	1213 (73.65%)	1647
<i>Hombre</i>	668 (24.16%)	2097 (75.84%)	2765
Tercer Quintil	1174 (26.61%)	3238 (73.39%)	4412
<i>Mujer</i>	612 (28.47%)	1538 (71.53%)	2150
<i>Hombre</i>	562 (24.85%)	1700 (75.15%)	2262
Cuarto Quintil	1079 (24.46%)	3333 (75.54%)	4412
<i>Mujer</i>	414 (26.01%)	1178 (73.99%)	1592
<i>Hombre</i>	665 (23.58%)	2155 (76.42%)	2820
Quinto Quintil	1113 (25.22%)	3300 (74.78%)	4413
<i>Mujer</i>	456 (24.68%)	1266 (73.52%)	1722
<i>Hombre</i>	657 (24.41%)	2034 (75.59%)	2691

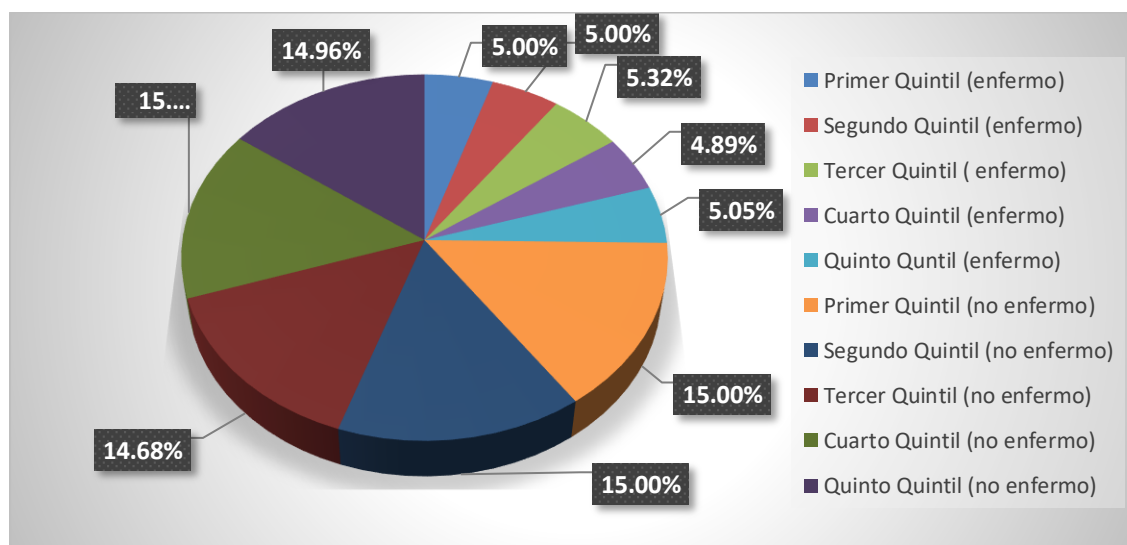
*La muestra incluye trabajadores entre 18 y 70 años.

Fuente: INEI (2019). *Encuesta Nacional de Hogares (ENAHOG)*. Recuperado el 24 de septiembre de 2020, de Instituto Nacional de Estadística e Informática: <http://inei.inei.gob.pe/microdatos/>. Elaboración propia.

La figura 3.6 muestra que la proporción de individuos enfermos pertenecientes a la desagregación de quintiles de ingreso representan alrededor del 25% del total. Lo anterior mencionado quiere decir que la cuarta parte de los individuos que fueron encuestados tienen la condición de haber estado enfermo en las últimas 4 semanas. Ahora, con respecto al total, la cantidad de encuestados enfermos con menor participación son los que pertenecen al cuarto quintil con 4.89% y los de mayor participación son los que pertenecen al quinto quintil con 5.05%. Finalmente, en cuanto a la cantidad de encuestados no enfermos, el quintil de ingreso con mayor participación con respecto al total es el cuarto quintil con 15,11% y el menor es el tercer quintil con 14.68%.

Figura 3.6

Número de individuos enfermos y no enfermos desagregado por quintiles de ingreso



Fuente: INEI (2019). *Encuesta Nacional de Hogares (ENAHOG)*. Recuperado el 24 de septiembre de 2020, de Instituto Nacional de Estadística e Informática: <http://inec.inei.gob.pe/microdatos/>. Elaboración propia.

La tabla 3.10 nos muestra el ingreso promedio desagregado por rangos de edad, sexo y el indicador de salud. Observamos con respecto a ella que las mujeres entre 18 y 24 años que tuvieron alguna enfermedad en las últimas 4 semanas tienen un ingreso promedio menor a las mujeres de ese mismo rango de edad que no tuvieron alguna enfermedad en las últimas 4 semanas. Las primeras contaban con un ingreso promedio de 551.95 soles y las segundas con 606.87 soles. En ese mismo sentido, las mujeres que están entre los 25 y 34 años y que tuvieron alguna enfermedad en las últimas 4 semanas tienen un ingreso promedio menor a las mujeres que no tuvieron alguna enfermedad en las últimas 4 semanas. Con relación a ese caso, los ingresos respectivos fueron de S/ 957.34 y S/ 1 044.91. Destacamos esos dos casos debido a que salieron significativos al 5% de significancia mediante la prueba de diferencia de medias que realizamos. En la tabla también se destaca también que el ingreso promedio de las personas entre 18 y 24 años (S/ 597.67) son menores que las personas que se encuentran entre los 25 y 34 años (S/ 1 099.82) y también que las personas que están entre los 35 y 44 años (S/ 1 206.41) y que las que están entre los 45 y 59 años (S/ 1 357.19). Lo anterior mencionado nos muestra que hay una secuencia en donde el ingreso promedio aumenta al aumentar de rango de edad, por lo que este último podría ser considerado un factor significativo para el ingreso promedio.

Tabla 3.10

Ingreso promedio desagregado por rangos de edad, indicador de salud y sexo

Rangos de edad	Estuvo enfermo en las últimas 4 semanas	No estuvo enfermo en las últimas 4 semanas	Total General	Test de Diferencia de Medias
18 -24	586.868	601.404	597.697	0.728
<i>Mujer</i>	551.953	606.874	591.024	1.952 **
<i>Hombre</i>	615.439	598.055	602.103	0.625
25 - 34	1074.047	1109.024	1099.820	1.081
<i>Mujer</i>	957.347	1044.914	1019.614	1.98 **
<i>Hombre</i>	1174.281	1152.859	1158.095	0.466
35 - 44	1175.034	1216.818	1206.409	1.027
<i>Mujer</i>	1032.766	1041.144	1039.060	0.156
<i>Hombre</i>	1288.165	1356.961	1339.805	1.172
45 - 59	1357.605	1357.055	1357.192	0.014
<i>Mujer</i>	1328.083	1302.472	1309.359	0.438
<i>Hombre</i>	1383.150	1396.328	1393.242	0.236
60 - 70	1235.771	1337.559	1314.098	1.175
<i>Mujer</i>	1256.262	1315.501	1300.951	0.423
<i>Hombre</i>	1225.402	1347.428	1320.153	1.116

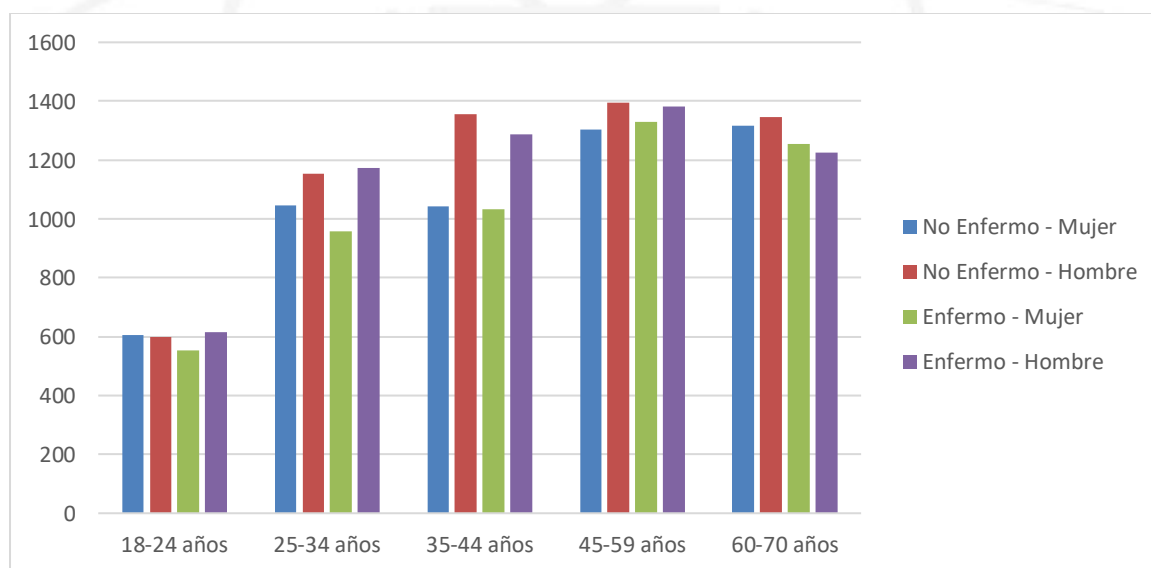
Fuente: INEI (2019). *Encuesta Nacional de Hogares (ENAHOG)*. Recuperado el 24 de septiembre de 2020, de Instituto Nacional de Estadística e Informática: <http://inei.inei.gob.pe/microdatos/>. Elaboración propia.

La figura 3.7 nos muestra que el ingreso promedio desagregado por rangos de edad, indicador de salud y enfermedad. Lo primero que destacamos es que los hombres enfermos como no enfermos tienen un ingreso promedio mayor tanto para las mujeres enfermas como no enfermas en todos los rangos de edad, excepto en el primer rango de edad, en donde las mujeres no enfermas tienen un ingreso promedio mayor que los hombres no enfermos.

Vemos también que el impacto del indicador de la salud es el esperado sobre los ingresos promedio para los hombres pertenecientes a los rangos de edad de 35 – 44, 45 – 59 y 60 – 70 años. Mientras que, para los hombres pertenecientes a los dos primeros rangos de edad, 18 – 24 y 25 – 34, no se cumple lo esperado. Por último, vemos que el impacto del indicador de salud es el esperado sobre los ingresos promedio para el caso de las mujeres en todos los rangos, excepto, para rango de 45 – 59 años.

Figura 3.7

Ingreso promedio desagregado por rangos de edad, indicador de salud y sexo



Fuente: INEI (2019). *Encuesta Nacional de Hogares (ENAH0)*. Recuperado el 24 de septiembre de 2020, de Instituto Nacional de Estadística e Informática: <http://iinei.inei.gob.pe/microdatos/>. Elaboración propia.

Por otra parte, la Tabla 3.11 muestra la división entre enfermos y no enfermos, subdividido por el género del entrevistado, que existe para cada rango de edad. En primer lugar, se puede observar que el rango de edad de 25 – 34 años es el que posee el mayor número de entrevistados con 5708 individuos. También verificamos que el tanto para los individuos no enfermos como para los enfermos, la proporción de hombres entrevistados

es mayor a la proporción de mujeres entrevistadas, asimismo, visualizamos que la cantidad de entrevistados no enfermos supera a la cantidad de entrevistados enfermos. Ahora, en cuanto al porcentaje relativo de enfermos con respecto a cada rango de edad, tenemos que la mayor proporción está en el rango de 25 – 34 años y la menor se encuentra en el rango de edad de 60 – 70 años. Por lo anterior mencionado se podría dar indicios de que los resultados econométricos sean más consistentes en el rango de 25 a 34 años en comparación a los demás rangos, dado que, tal como se mencionó al inicio de la descripción, cuenta con el mayor número de individuos entrevistados.

Tabla 3.11

Número de individuos enfermos y no enfermos desagregado por rangos de edad

Rangos de Edad	Número de individuos enfermos	Número de individuos no enfermos	Total de individuos
18 - 24 años	1051 (25.50%)	3070 (74.50%)	4121
<i>Mujer</i>	473 (28.86%)	1166 (71.14%)	1639
<i>Hombre</i>	578 (23.29%)	1904 (76.71%)	2482
25 - 34 años	1502 (26.31%)	4206 (73.69%)	5708
<i>Mujer</i>	694 (28.89%)	1708 (71.11%)	2402
<i>Hombre</i>	808 (24.44%)	2498 (75.56%)	3306
35 - 44 años	1271 (24.91%)	3831 (75.09%)	5102
<i>Mujer</i>	563 (24.88%)	1700 (75.12%)	2263
<i>Hombre</i>	708 (24.94%)	2131 (75.06%)	2839
45 - 59 años	1371 (24.91%)	4132 (75.09%)	5503
<i>Mujer</i>	636 (26.89%)	1729 (73.11%)	2365
<i>Hombre</i>	735 (23.42%)	2403 (76.58%)	3138
60 - 70 años	375 (23.05%)	1252 (76.95%)	1627
<i>Mujer</i>	126 (24.56%)	387 (75.44%)	513
<i>Hombre</i>	249 (22.35%)	865 (77.65%)	1114

La muestra incluye trabajadores entre 18 y 70 años.

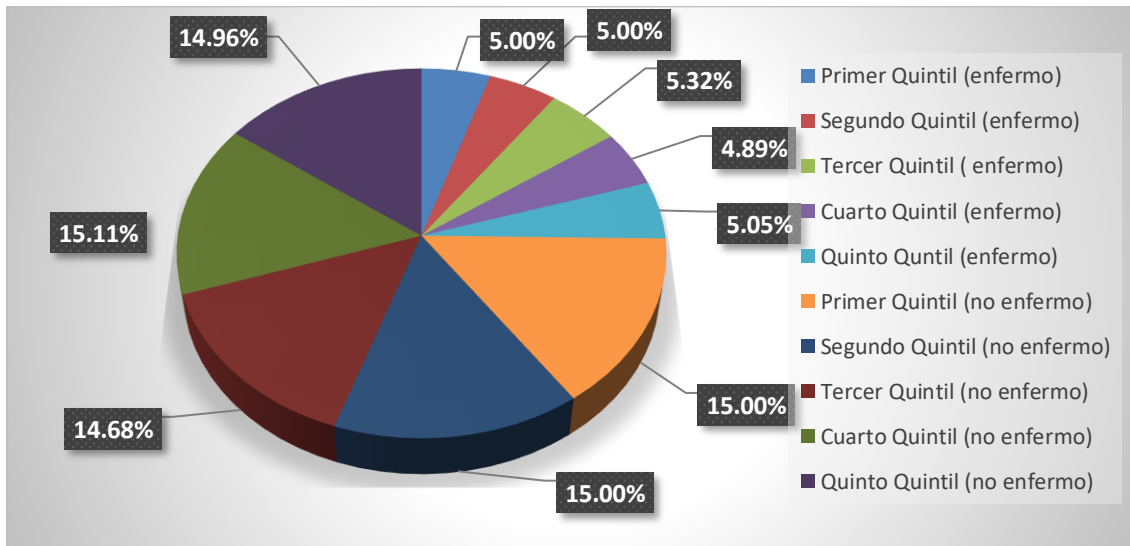
Fuente: INEI (2019). *Encuesta Nacional de Hogares (ENAHO)*. Recuperado el 24 de septiembre de 2020, de Instituto Nacional de Estadística e Informática: <http://inei.inei.gob.pe/microdatos/>. Elaboración propia.

La figura 3.8 nos muestra que la proporción de individuos enfermos representan el 25% del total, es decir, la cuarta parte de los individuos encuestados tienen la condición de haber estado enfermo en las últimas 4 semanas. Además, respecto al total, la cantidad de encuestados enfermos con menor participación es del nivel educativo primaria con 3%

y la de mayor participación es del nivel educativo superior con 12%. En cuanto a la cantidad de encuestados no enfermos, el nivel educativo con mayor participación es el superior con 34% y el nivel educativo con menor participación es el nivel primario con 11%.

Figura 3.8

Número de individuos enfermos y no enfermos desagregado por rangos de edad



Fuente: INEI (2019). *Encuesta Nacional de Hogares (ENAHO)*. Recuperado el 24 de septiembre de 2020, de Instituto Nacional de Estadística e Informática: <http://inei.inei.gob.pe/microdatos/>. Elaboración propia.

CAPÍTULO IV: ANÁLISIS DE RESULTADOS

A lo largo de la presente investigación hemos buscado encontrar el impacto que tiene la salud, variable del capital humano, sobre la productividad. En ese sentido, sobre la base de la Encuesta Nacional de Hogares del 2019 hemos podido hallar los resultados finales que se presentarán en esta sección. Para iniciar, la tabla 4.1 muestra el resultado final a nivel general, considerando a los 22,061 encuestados que cumplían con los requisitos para la investigación.

Tabla 4.1

Resultado econométrico a nivel general mediante el modelo MCO corregido por la estimación en dos etapas de Heckman

Variables	Regresión General
1 Constante	-244.087(*)
<i>Características Individuales</i>	
2 Edad	30.521(***)
3 Cuadrado de la Edad	-0.291(***)
4 Sexo	243.242(***)
<i>VARIABLES DE CAPITAL HUMANO</i>	
5 Años de Educación	-134.633(***)
6 Cuadrado de los años de Educación	12.848(***)
7 Años de Experiencia	25.876(***)
8 Cuadrado de los años de Experiencia	-0.176(**)
9 Indicador de Salud	29.645(**)
<i>VARIABLES SOBRE LA VIVIENDA</i>	
10 Baño con Desagüe	-34.137(***)
11 Horas de Oferta de Agua	6.454(***)
12 Teléfono Fijo	197.382(***)
13 Conexión a Internet	202.225(***)
<i>VARIABLES DEL MERCADO LOCAL</i>	
14 Total de horas trabajadas	8.192(***)
15 Término de selección	-2166.520(***)
<i>Tasa de retorno de la educación</i>	12.120%
<i>Tasa de retorno del indicador de salud</i>	2.670%
R Cuadrado Ajustado	0.3973
Número de observaciones	22061

(*) Significancia al 10% de confianza; (**) Significancia al 5% de confianza; (***) significancia al 1% de confianza.

Fuente: INEI (2019). *Encuesta Nacional de Hogares (ENAHOG)*. Recuperado el 24 de septiembre de 2020, de Instituto Nacional de Estadística e Informática: <http://inei.inei.gob.pe/microdatos/>. Elaboración propia.

La tabla 4.1 muestra los resultados obtenidos por el modelo de Heckman, ello nos indica un resultado esperado de acuerdo con la variable número 9 que es el indicador de salud. Este resultado nos indica que el no haber estado enfermo en las últimas 4 semanas impacta positivamente sobre los ingresos salariales y ese resultado es significativo al 5%; es decir, se comprueba que la variable salud tiene un impacto positivo y significativo durante el 2019. Con base en el valor reportado por el coeficiente del indicador de salud, se halló la tasa de retorno de la salud que resultó en 2.67% con respecto al ingreso promedio. Lo anterior es consistente con el resultado hallado por Cortez (1999), Ribero (2000), Schultz y Tansel (1997), entre otros autores mencionados en la literatura empírica, que encuentran un impacto significativo y positivo del estado de salud sobre la productividad.

Cabe precisar que, respecto a las características individuales, la edad muestra el comportamiento positivo esperado, y el cuadrado de la edad es consistente con el carácter decreciente de la variable edad. El sexo es una variable significativa y nos muestra que los hombres en promedio ganan 243 soles más que las mujeres, lo cual es consistente con la brecha salarial que existe actualmente en el Perú.

Por otro lado, las variables que completan el capital humano nos muestra un comportamiento esperado en los años de experiencia, puesto que su signo es positivo y el cuadrado de la experiencia nos muestra la no linealidad de la misma variable; ahora bien, en cuanto a la variable años de educación vemos que no se cumple lo predicho por la teoría y el cuadrado de los años de educación confirma un comportamiento creciente de los años de educación, ello puede explicarse probablemente a que los frutos de la educación recién se ven a partir de la educación superior en adelante. Además, la tasa de retorno de la educación calculada es de 12.12% con respecto al ingreso promedio.

En cuanto a las variables sobre la vivienda, vemos que el acceso a servicios de internet, agua potable y teléfono fijo impactan positivamente sobre los ingresos salariales, mientras que el acceso a un sistema de desagüe adecuado tiene un comportamiento contrario al esperado. Cabe señalar que los coeficientes de todas las variables antes mencionadas presentan una significancia al 1%. Por último, tenemos la variable de total de horas trabajadas que también corresponde lógicamente con un signo positivo y significativo, esto puede interpretarse de tal forma que por una hora más de trabajo se gana aproximadamente 8.2 soles más.

Ahora bien, con respecto a las 4 desagregaciones que planteamos en el estudio, la tabla 4.2 muestra los resultados con respecto a la desagregación de los 3 niveles educativos considerados, estos son: primaria, secundaria y superior.

Tabla 4.2

Resultado econométrico sobre los niveles educativos mediante el modelo MCO corregido por la estimación en dos etapas de Heckman

Variables	Niveles Educativos		
	Primaria	Secundaria	Superior
1 Constante	-51.06	-483.54	20901.98(***)
<i>Características Individuales</i>			
2 Edad	5.03	6.68	27.83(**)
3 Cuadrado de la Edad	-0.02	-0.04	-0.21
4 Sexo	125.92(***)	246.73(***)	280.98(***)
<i>VARIABLES DE CAPITAL HUMANO</i>			
5 Años de Educación	16.73	80.94	-3113.89(***)
6 Cuadrado de los años de Educación	-1.92	-1.17	116.47(***)
7 Años de Experiencia	-2.87	2.25	51.55(***)
8 Cuadrado de los años de Experiencia	0.23(*)	0.57(***)	-0.80(***)
9 Indicador de Salud	-23.29	3.37	59.82(**)
<i>VARIABLES SOBRE LA VIVIENDA</i>			
10 Baño con Desagüe	47.69(***)	45.14(***)	-42.43
11 Horas de Oferta de Agua	2.14(**)	3.90(***)	7.81(***)
12 Teléfono Fijo	31.60	109.53(***)	180.96(***)
13 Conexión a Internet	124.11(***)	184.68(***)	284.16(***)
<i>VARIABLES DEL MERCADO LOCAL</i>			
14 Total de horas trabajadas	8.74(***)	7.07(***)	10.71(***)
15 Término de selección	-638.7(***)	-1591.78(***)	-3454.27(***)
<i>Tasa de retorno de la educación</i>	3.691%	10.909%	19.176%
<i>Tasa de retorno del indicador de salud</i>	5.138%	0.454%	3.653%
R Cuadrado Ajustado	0.1726	0.2142	0.3317
Número de observaciones	3142	8822	10097

(*) Significancia al 10% de confianza; (**) Significancia al 5% de confianza; (***) significancia al 1% de confianza.

Fuente: INEI (2019). *Encuesta Nacional de Hogares (ENAHO)*. Recuperado el 24 de septiembre de 2020, de Instituto Nacional de Estadística e Informática: <http://inei.inei.gob.pe/microdatos/>. Elaboración propia.

En relación con la primera desagregación que planteamos sobre los 3 niveles educativos, luego de realizar la regresión, encontramos que el indicador de salud es significativo y con el signo positivo esperado para el nivel superior. En ese sentido, podemos interpretar que una persona con estudios superiores que no ha presentado

enfermedad en las últimas 4 semanas tiene un ingreso promedio mayor en 59.8 soles con respecto a una persona del mismo nivel educativo que sí ha estado enfermo. Además, se establece que la tasa de retorno de la salud para ese nivel educativo es de 3.65% respecto al ingreso promedio.

Podemos notar que para los niveles secundaria y primaria, el resultado es no significativo, también podemos visualizar cómo ha evolucionado el indicador de salud pasando de tener un signo no esperado y no significativo en el nivel primaria a tener un signo esperado y no significativo en el nivel secundaria; y que finalmente, es en el nivel superior donde se ve que es significativo al 5% y consistente con el signo esperado. Al comienzo de la investigación, habíamos planteado como hipótesis que el impacto de la salud sobre la productividad iba a ser mayor para los niveles educativos más básicos como el nivel primario, seguido del nivel secundaria y finalmente el superior. Ello porque se espera que las personas con menos estudios se dediquen a labores que demandan más esfuerzo físico y mantener una buena salud sea primordial para su trabajo y salario, por lo tanto, la hipótesis planteada para esta desagregación resulta contraria a los resultados encontrados con la econometría. Ahora bien, lo anterior se vislumbraba desde las estadísticas descriptivas mostradas sobre la desagregación de los niveles educativos, ya que la prueba de diferencia de medias solo era significativa para el nivel superior y al 1% de significancia; otra posible razón es que se le esté restando importancia a la variable salud en los niveles más básicos y recién cuando una persona alcanza el nivel superior toma mayores medidas para la protección de su salud y que ello puede estar reflejándose en los resultados.

En cuanto a las características individuales, la variable sexo es significativa al 1% para los 3 niveles educativos considerados, lo cual refuerza la importancia de esta variable sobre los ingresos salariales. En cuanto a la variable de la edad y el cuadrado de la edad, los signos indican un comportamiento decreciente consistente con la teoría, pero la significancia de las variables solo es válida para el nivel superior al 5% de significancia. Además, con relación a las variables de capital humano, vemos que casi solo en el nivel superior se ve que sean significativas y con el signo esperado.

Respecto a las variables de la vivienda, la mayoría sigue siendo significativa y con el signo esperado a excepción del baño con desagüe en el nivel superior. Sobresale la variable de conexión a internet al ser significativa al 1% en los 3 niveles; ello puede

sostenerse sobre la base de que el acceso a internet es necesario para la educación y el acceso a mejores oportunidades laborales que impactan sobre los ingresos. %. Por último, tenemos la variable de total de horas trabajadas que también corresponde lógicamente con un signo positivo y significativo, esto puede interpretarse de tal forma que por una hora más de trabajo se gana aproximadamente 10.71 soles más en el nivel superior, 8.74 soles más en el nivel primaria y 7.07 soles más en el nivel secundaria; ello resulta consistente con el hecho de que en el nivel superior se obtenga una mayor recompensa por las horas trabajadas.

La tabla 4.3 muestra los resultados con respecto a la desagregación de los 8 dominios geográficos. En relación con esta desagregación que planteamos sobre los 8 dominios geográficos, encontramos que, luego de realizar la regresión, obtenemos que el indicador de salud muestra un comportamiento significativo y positivo en los dominios de Lima Metropolitana al 1% de significancia y en la Sierra Sur al 10% de significancia, un comportamiento negativo y no esperado en la Sierra Norte al 5% de significancia y para los demás dominios no es significativo. En ese sentido podemos establecer las tasas de retorno para verificar el impacto del indicador de salud, verificamos que el mayor impacto se da en los dominios de Lima Metropolitana con una tasa de 11.59% y en Sierra Norte con 12.69%; y en los que se muestra menor impacto es en los dominios de la Selva con 0.75% y la Costa Norte con 0.85%. Entonces, vemos que la hipótesis esperada no se cumple y, por el contrario, el dominio de Lima Metropolitana que esperábamos que iba a estar entre los que la variable de Salud impacta en menor proporción que los demás, resulta opuesto a lo que encontramos, ya que justamente en Lima Metropolitana el impacto es de los más altos, con el signo esperado y a un nivel de significancia del 1%.

Cabe señalar que los resultados de la tabla 4.3, con relación al indicador de salud, se esperaban dado que la prueba de diferencia de medias indicaba una diferencia significativa solo para los dominios de Lima Metropolitana al 1%, Sierra Norte al 5% y Sierra Sur al 5%; por tanto, la econometría confirma lo que esperábamos con las estadísticas descriptivas iniciales. El caso particular de la Sierra Norte también se visualizó en la tabla 3.4.3 donde las personas que estaban enfermas ganaban más que las personas sanas, además, esa diferencia era significativa al 5%. Cabe señalar que una posible razón sea la muestra reducida de las personas enfermas en la Sierra Norte, ya que representan solo el 1.21% respecto del total, siendo la muestra más pequeña.

Tabla 4.3

Resultado econométrico sobre los dominios geográficos mediante el modelo MCO corregido por la estimación en dos etapas de Heckman

Variables	Dominios geográficos							
	Costa Norte	Costa Centro	Costa Sur	Lima Metropolitana	Selva	Sierra Norte	Sierra Centro	Sierra Sur
1 Constante	-466.01	-310.92	-2030.36(***)	-3276.75(***)	-879.53(**)	-356.18	-1359.36(***)	459.13
<i>Características Individuales</i>								
2 Edad	30.58(***)	29.08(***)	83.21(***)	126.66(***)	44.23(***)	40.25(***)	65.78(***)	3.16
3 Cuadrado de la Edad	-0.28(**)	-0.28(***)	-0.92(***)	-1.44(***)	-0.4(***)	-0.33(**)	-0.72 (***)	0.02
4 Sexo	135.01(***)	130.61(***)	468.78(***)	265.34(***)	158.01(***)	137.7(***)	275.84(***)	437.39(***)
<i>Variables de capital humano</i>								
5 Años de Educación	109.69(***)	-125.11(***)	-88.89(***)	-249.78(***)	-94.01(***)	-95.53(***)	-74.49(***)	-150.77(***)
6 Cuadrado de los años de Educación	11.27(***)	11.19(***)	10.22(***)	21.54(***)	11.67(***)	10.31(***)	9.05(***)	12.75(***)
7 Años de Experiencia	24.11(***)	30.01(***)	33.75(***)	45.2(***)	20.55(***)	29.8(***)	28.19(***)	26.58(***)
8 Cuadrado de los años de Experiencia	-0.27(*)	-0.39(***)	0.08	-0.64(***)	-0.25(*)	-0.56(**)	-0.28	0.21
9 Indicador de Salud	7.73	14.79	-67.80	172.14(***)	7.61	-134.45(**)	18.35	69.15(*)
<i>Variables sobre la vivienda</i>								
10 Baño con Desagüe	69.69(***)	-11.57	-34.65	-130.21(**)	28.07	-12.71	52.31(*)	15.46
11 Horas de Oferta de Agua	-1.82	3.98(***)	0.35	16.66(***)	0.00	-7.17(***)	1.90	3.31(*)
12 Teléfono Fijo	53.32	252.60(***)	270.05(***)	249.44(***)	225.1(***)	454.31(**)	62.77	-54.19

Variables	Dominios geográficos							
	Costa Norte	Costa Centro	Costa Sur	Lima Metropolitana	Selva	Sierra Norte	Sierra Centro	Sierra Sur
13 Conexión a Internet	208.07(***)	158.90(***)	218.51(***)	236.96(***)	182.62(***)	200.46(***)	261.32(***)	322.04(***)
<i>Variables del mercado local</i>								
14 Total de horas trabajadas	7.16(***)	7.14(***)	11.13(***)	7.45(***)	6.14(***)	8.14(***)	8.36(***)	13.04(***)
15 Término de selección	-1329.6(**)	-1613.31(***)	129.43	3636.74(***)	-812.88	-1337.44(***)	-581.71(**)	-3021.35(***)
<i>Tasa de retorno de la educación</i>	12.103%	13.902%	7.294%	16.82%	9.21%	9.02%	7.10%	11.77%
<i>Tasa de retorno del indicador de salud</i>	0.852%	1.643%	5.563%	11.59%	0.745%	12.695%	1.750%	5.397%
R Cuadrado Ajustado	0.4089	0.4235	0.4011	0.3287	0.5231	0.5392	0.4713	0.4137
Número de observaciones	3972	2949	1661	3733	3785	956	2621	2375

(*) Significancia al 10% de confianza; (**) Significancia al 5% de confianza; (***) significancia al 1% de confianza.

Fuente: INEI (2019). *Encuesta Nacional de Hogares (ENAH0)*. Recuperado el 24 de septiembre de 2020, de Instituto Nacional de Estadística e Informática: <http://inei.inei.gob.pe/microdatos/>. Elaboración propia.

Respecto a las demás variables, en la tabla 4.3, vemos que hay un comportamiento esperado en la edad y cuadrado de la edad a un nivel de significancia del 1% en todos los dominios, excepto en el dominio de Sierra Sur. Vemos también que la variable Sexo es significativa al 1% en todos los dominios, y el mayor impacto se da en la Costa Sur. En cuanto a las variables de capital humano, los años de educación y cuadro de años de educación tienen el comportamiento creciente que vimos anteriormente, y los años de experiencia junto con el cuadrado de los años de experiencia muestran los signos esperados y significativos en la mayoría de los dominios.

Sobre las variables de la vivienda vemos que el comportamiento de las variables es desigual, pero en la variable de conexión a internet vemos que es significativa al 1% en todos los dominios y con el signo positivo esperado, el mayor impacto se da en la zona de la Sierra Sur. Por último, tenemos la variable de total de horas trabajadas que también corresponde lógicamente con un signo positivo y significativo; respecto a esta variable, vemos que el dominio de la Costa Sur tiene un coeficiente de 11.13 y la Sierra Sur tiene un coeficiente de 13.04, es decir, en promedio se paga más por hora trabajada en la zona sur del país. Por el contrario, en el dominio de la selva se paga menos por hora trabajada, esto es, alrededor de 6.14 soles más por cada hora trabajada.

La tabla 4.4 muestra los resultados en relación con la tercera desagregación que planteamos sobre los quintiles de ingreso, en ese sentido, encontramos que el indicador de salud es significativo y con el signo positivo esperado para el segundo, cuarto y quinto quintil. Se puede interpretar así que los individuos pertenecientes al segundo, cuarto y quinto quintil que no han presentado alguna enfermedad en las últimas 4 semanas tienen un ingreso promedio mayor de 6.8, 14.97 y 107,14 respectivamente con respecto a sus pares en esos quintiles que sí presentaron alguna enfermedad en las últimas 4 semanas.

El mayor impacto del indicador de salud se puede verificar gracias a la tasa de retorno de la salud, con ello confirmamos que el mayor impacto se da en el quinto quintil con una tasa de 3.7% al 1% de significancia, le sigue el segundo quintil con una tasa de 1.9% al 5% de significancia y por último el cuarto quintil con una tasa de 1.1% al 5% de significancia, todas ellas con respecto al ingreso promedio.

Tabla 4.4

Resultado econométrico sobre los quintiles de ingreso mediante el modelo MCO corregido por la estimación en dos etapas de Heckman

Variables	Quintiles de Ingreso				
	Quintil 1	Quintil 2	Quintil 3	Quintil 4	Quintil 5
1 Constante	76.97(***)	357.54(***)	655.85(***)	1076.89(***)	711.98(**)
<i>Características Individuales</i>					
2 Edad	0.39	-0.53	-0.21	7.35(***)	85.59(***)
3 Cuadrado de la Edad	0.00	0.01	0.00	-0.08(***)	-0.68(***)
4 Sexo	17.28(***)	-18.36(***)	4.87	67.68(***)	336.12(***)
<i>VARIABLES DE CAPITAL HUMANO</i>					
5 Años de Educación	-0.64	-1.21	-0.44	-9.54(***)	-223.36(***)
6 Cuadrado de los años de Educación	0.11	0.18	0.15	1.2(***)	14.78(***)
7 Años de Experiencia	0.59(**)	-0.73	2.12(**)	2.84(**)	2.94
8 Cuadrado de los años de Experiencia	-0.01	0.01	-0.05(*)	-0.01	-0.18
9 Indicador de Salud	0.89	6.8(**)	-2.53	14.97(**)	107.14(***)
<i>VARIABLES SOBRE LA VIVIENDA</i>					
10 Baño con Desagüe	-0.21	11.1(***)	15.46(**)	-4.98	-42.86
11 Horas de Oferta de Agua	-0.01	0.38(**)	1.17(***)	0.57	9.45(***)
12 Teléfono Fijo	2.54	8.85(*)	19.4(***)	8.64	287.61(***)
13 Conexión a Internet	7.12(***)	1.36	7.39	-12.09	207.92(***)
<i>VARIABLES DEL MERCADO LOCAL</i>					

Variables	Quintiles de Ingreso				
	Quintil 1	Quintil 2	Quintil 3	Quintil 4	Quintil 5
14 Total de horas trabajadas	1.41(***)	0.31(***)	1.21(***)	0.21	9.02(***)
15 Término de selección	-49.2(*)	8.80	568.43(**)	-174.66	-6455.91(***)
<i>Tasa de retorno de la educación</i>	0.529%	0.348%	0.053%	0.698%	7.722%
<i>Tasa de retorno del indicador de salud</i>	0.735%	1.955%	0.306%	1.095%	3.704%
R Cuadrado Ajustado	0.2387	0.045	0.0633	0.0679	0.1248
Número de observaciones	4,412	4,412	4,412	4,412	4,413

(*) Significancia al 10% de confianza; (**) Significancia al 5% de confianza; (***) significancia al 1% de confianza.

Fuente: INEI (2019). *Encuesta Nacional de Hogares (ENAHO)*. Recuperado el 24 de septiembre de 2020, de Instituto Nacional de Estadística e Informática: <http://inei.inei.gob.pe/microdatos/>. Elaboración propia.

Respecto a los quintiles de ingreso, habíamos planteado como hipótesis específica que el impacto de la salud sobre la productividad iba a ser mayor para los quintiles más pobres, esto era porque se esperaba que, al tener menor presupuesto para invertir en salud, estos estaban más propensos a estar enfermos y padecer más dolencias que afectarían su productividad. Respecto a lo anterior, los resultados econométricos no concuerdan totalmente, ya que el mayor impacto se presenta en el quinto quintil, pero sí concuerda parcialmente toda vez que se encuentra impacto en uno de los rangos más pobres, como lo es el segundo quintil. Además, se observa que, para el primer y tercer quintil, el resultado es no significativo; ahora bien, lo anterior se esperaba a partir de las estadísticas descriptivas de los quintiles de ingreso, dado que la prueba de diferencia de medias era significativa para todos los quintiles, excepto para el tercer quintil. Para explicar la no significancia del primer quintil, la razón sería que se le está restando importancia a la variable salud para darle mayor relevancia a la variable años de experiencia como factor principal para elevar la productividad. En el quinto quintil, por el contrario, sí se ve el mayor impacto del indicador de salud, por lo que se podría afirmar que se toman mayores medidas para la protección de la salud, ya que al tener mayores ingresos tienen mayor renta disponible para invertir en salud que puede generar un círculo virtuoso con la productividad, lo cual puede estar reflejándose en los resultados.

En cuanto a las características individuales, la variable sexo es significativa al 1% para 4 de los 5 quintiles considerados, la excepción se da en el tercer quintil, en donde se toma más en cuenta a variables del capital humano y de la vivienda a la hora de elevar la productividad. De igual manera destaca esta variable por su implicancia sobre los ingresos salariales de los otros quintiles con una significancia al 1%. Con relación a las variables de capital humano, vemos que la variable años de experiencia es relevante para casi todos los quintiles, excepto el segundo quintil, donde es no significativo.

Por último, con respecto a las variables sobre la vivienda se tiene a la variable de total de horas trabajadas, que es significativa para todos los quintiles, excepto para el cuarto quintil. Esto puede interpretarse de tal forma que por una hora más de trabajo se gana aproximadamente 9.02 soles más en el quinto quintil, 1.41 soles más en el primer quintil, 1.21 soles más en el tercer quintil y se gana 0.31 soles en el segundo quintil, todo ello resulta consistente con el hecho de que en el quinto quintil se obtenga una mayor recompensa por las horas trabajadas.

Tabla 4.5

Resultado econométrico sobre los rangos de edad mediante el modelo MCO corregido por la estimación en dos etapas de Heckman

Variables	Rangos de Edad				
	18-24	25-34	35-44	45-59	60-70
1 Constante	3054.99(***)	3147.69(**)	-4317.52	864.40	-3776.57
<i>Características Individuales</i>					
2 Edad	-219.41(**)	226.86(**)	255.06(*)	5.90	117.28
3 Cuadrado de la Edad	4.58(**)	-3.21(*)	-3.06(*)	-0.16	-0.90
4 Sexo	64.58(***)	227.56(***)	382.27(***)	276.9(***)	305.76(***)
<i>VARIABLES DE CAPITAL HUMANO</i>					
5 Años de Educación	-55.59(***)	243.87(***)	-291.81(***)	-171.46(***)	-167.3(***)
6 Cuadrado de los años de Educación	5.32(***)	16.79(***)	20.43(***)	15.93(***)	16.37(***)
7 Años de Experiencia	30.36(***)	75.65(***)	52.98(***)	32.3(***)	9.38
8 Cuadrado de los años de Experiencia	-5.47(***)	-5.03(***)	-2.08(***)	-0.5(***)	0.33(*)
9 Indicador de Salud	23.69	36.91	64.64(**)	-4.44	11.81
<i>VARIABLES SOBRE LA VIVIENDA</i>					
10 Baño con Desagüe	-34.68(**)	40.53(*)	9.00	-38.06	108.23(**)
11 Horas de Oferta de Agua	4.39(***)	6.21(***)	7.24(***)	4.81(***)	8.61(***)
12 Teléfono Fijo	60.18(**)	107.15(**)	240.32(***)	257.73(***)	231.39(***)
13 Conexión a Internet	99.66(***)	227.59(***)	307.91(***)	208.12(***)	254.98(***)

Variables	Rangos de Edad				
	18-24	25-34	35-44	45-59	60-70
<i>Variables del mercado local</i>					
14 Total de horas trabajadas	9.06(***)	8.38(***)	7.25(***)	6.66(***)	10.18(***)
15 Término de selección	-1673.45(***)	949.16(***)	-1697.57(***)	-2717.02(***)	-1886.47(***)
<i>Tasa de retorno de la educación</i>	9.301%	22.174%	24.188%	12.633%	12.731%
<i>Tasa de retorno del indicador de salud</i>	3.964%	3.356%	5.358%	0.327%	0.899%
R Cuadrado Ajustado	0.2558	0.3389	0.4029	0.4295	0.4187
Número de observaciones	4121	5708	5102	5503	1627

(*) Significancia al 10% de confianza; (**) Significancia al 5% de confianza; (***) significancia al 1% de confianza.

Fuente: INEI (2019). *Encuesta Nacional de Hogares (ENAHO)*. Recuperado el 24 de septiembre de 2020, de Instituto Nacional de Estadística e Informática: <http://inei.inei.gob.pe/microdatos/>. Elaboración propia.

Por último, la tabla 4.5 muestra los resultados con relación a la cuarta desagregación y esperábamos que la diferencia del impacto de la salud sobre la productividad entre los cinco rangos de edad fuera mayor para los rangos de edad más jóvenes, debido a que, generalmente, estos son los que requieren realizar un mayor esfuerzo físico en sus labores, en comparación con los rangos de edad más adultos, que habitualmente se dedican a realizar labores que demandan más experiencia.

Al realizar la regresión, encontramos que el indicador de salud muestra un comportamiento positivo y significativo únicamente en el tercer rango de edad, 35 a 44 años, al 5% de significancia. Luego de ello también podemos verificar el impacto del indicador de salud a través de la tasa de retorno de esta, y verificamos que el mayor impacto se da en el tercer rango de edad con una tasa de 5.36%, le sigue el primer rango de edad con una tasa de 3.96% y luego el segundo rango con una tasa de 3.36%; ahora en los que se muestra un menor impacto son el cuarto y quinto rango de edad con tasas de 0.33% y 0.90% respectivamente. Es así que vemos que la hipótesis esperada no se cumple del todo, pero sí parcialmente.; afirmamos ello debido a que el mayor impacto de la salud sobre la productividad se da en el tercer rango de edad y no en los rangos de edad más jóvenes, entre el primero y segundo. Esto último se podría explicar debido a que en el tercer rango de edad es donde se llega a la madurez laboral y, por tanto, se logra estabilidad monetaria que permite gastar más en servicios de salud en comparación a los rangos de edad más jóvenes.

En cuanto a las demás variables, vemos que tienen un comportamiento esperado en la edad y cuadrado de la edad a un nivel de significancia entre 5% y 10% en todos los dominios, excepto en los dos últimos rangos de edad en donde no son significativos. Vemos también que la variable sexo es significativa al 1% en todos los dominios, lo cual termina por reforzar la gran relevancia del sexo sobre los ingresos. En cuanto a las variables de capital humano, vemos que los años de educación y cuadro de años de educación son significativas al 1% de significancia en todos los rangos de edad, y los años de experiencia junto con el cuadrado de los años de experiencia muestran significancia en todos los rangos, excepto en el quinto rango de edad.

Sobre las variables de la vivienda vemos que el comportamiento de las variables es muy homogéneo, las variables horas de oferta de agua, acceso a teléfono fijo, conexión a internet y total de horas trabajadas tienen el signo positivo esperado y la significancia

correspondiente. Ahora, en cuanto a la variable baño con desagüe, esta es significativa en el primer, segundo y quinto rango de edad; mientras que en el tercer y cuarto rango de edad no hay significancia en esta variable. Cabe resaltar que la variable mencionada anteriormente es la única que no es significativa para el tercer rango de edad, en donde el indicador de salud únicamente es positivo y significativo. Concluimos esta parte afirmando que es en el quinto rango de edad en donde se paga en promedio más, por hora trabajada, con 10.18 soles; y, en el cuarto rango de edad en donde se paga menos por hora trabajada, con 6.66 soles.



CAPÍTULO V: CONCLUSIONES

- En la presente investigación se analizó el efecto de la salud, como variable del capital humano, sobre la productividad que fue medida en términos de los ingresos mensuales de los encuestados considerados en la investigación; para cumplir con tal fin, hemos empleado el modelo en dos etapas de Heckman que incorpora una corrección al error de autoselección. Los resultados generales sugieren que hay un impacto significativo y positivo de la salud sobre la productividad, ello reafirma lo planteado en la hipótesis general, lo cual es consistente con los resultados encontrados por autores como Cortez (1999), Ribero (2000), Thomas y Strauss (1997), Schultz y Tansel (1997) y otros autores, quienes hallaron un impacto significativo y positivo de la salud sobre la productividad en sus respectivos trabajos de investigación.
- Las desagregaciones planteadas en la investigación muestran resultados mixtos respecto a lo esperado, sobre los resultados por dominios geográficos, se encontró que el mayor efecto significativo y positivo se da sobre el dominio de Lima Metropolitana, que tiene una tasa de retorno de la salud de 11.59% y significativa al 1%. Este resultado contrasta con la hipótesis planteada inicialmente, ya que se esperaba que el impacto de la salud fuera mayor en dominios menos industrializados, como la Selva o algunas zonas de la Sierra. En ese sentido, podemos determinar que existe un rol principal de la salud en Lima Metropolitana y Sierra Sur, y un papel secundario en dominios como la Selva, Sierra Centro y la Costa, en estos últimos dominios es probable que se esté dando mayor importancia a variables como los años de educación y la experiencia que también son significativos y positivos en esos dominios.
- Los resultados por niveles educativos muestran que solamente en el nivel superior se encuentra un efecto significativo y positivo de la salud sobre la productividad, donde la tasa de retorno calculada es de 3.65% y significativa al 5%. Este resultado contrasta con la hipótesis y sugiere que, cuando se alcanza el nivel de educación superior, se otorga mayor relevancia a los beneficios de la salud sobre la productividad.
- La desagregación por quintiles de ingreso muestra un impacto positivo y significativo de la salud para el segundo, cuarto y quinto quintil. La tasa de retorno de la salud más

alta se verifica en el quinto quintil, lo cual contrasta con la hipótesis, ello implica que aquellos que disponen de mayores ingresos pueden verse más afectados ante un evento que impacte negativamente sobre la salud, en comparación con aquellos que disponen de menores ingresos; lo cual es consistente con el hecho que aquellos que disponen de mayores ingresos pueden destinar más recursos destinados al cuidado de la salud.

- Finalmente, con relación a la desagregación por rangos de edad, encontramos que el mayor efecto se da únicamente en el tercer rango de edad, 35 – 44 años, con una significancia del 5% y una tasa de retorno de la salud de 5.36%. Ello contrasta con la hipótesis, y ello se puede deber a que generalmente en este rango de edad, las personas llegan a una madurez laboral que les permite realizar mayores gastos de salud.



CAPÍTULO VI: RECOMENDACIONES

- Dado que, la medición cualitativa del indicador de salud genera mayor complejidad en el análisis de impacto; así como, resultados menos precisos en las regresiones econométricas. Se recomienda a las instituciones encargadas de realizar las encuestas y recopilar la información estadística correspondiente, como el Instituto Nacional de Estadística e Informática (INEI), incorporar métodos que permitan la medición del indicador de salud a través de variables cuantitativas, ello permitiría resultados econométricos más precisos y con mayor valor para las interpretaciones económicas. Además, es recomendable una correcta formulación de preguntas hacia los encuestados con relación al estado de salud para un posterior análisis, esto debido a que muchas de las respuestas se basan en la percepción personal de lo que significa estar enfermo y se podrían presentar problemas de heterocedasticidad. Para ello recomendamos que haya una corroboración de como afectó el estar enfermo mediante preguntas complementarias.
- El estado cumple un rol fundamental para la reducción de las brechas de infraestructura o de acceso a servicios; en ese sentido, dada la importancia de reducir las brechas existentes en el sector salud; se recomienda a las entidades públicas correspondientes (Ministerio de Salud – MINSA, EsSalud, Sanidad de las Fuerzas Armadas, Sanidad de las Fuerzas Policiales, etc.), que para el cumplimiento de los objetivos planteados en las políticas públicas, formulen sus documentos de gestión pública tomando en consideración que el impacto de la inversión en salud no es igual para todos. En efecto, se recomienda tomar en consideración que las diversas intervenciones públicas, como la realización de campañas informativas sobre de los beneficios de una buena salud, tendrán mayor relevancia en los dominios geográficos, rangos de edad, quintiles de ingresos y nivel educativo donde la variable de la salud no alcanza suficiente significancia.
- En línea con lo anterior, recomendamos la priorización de proyectos de inversión destinados a la creación de centros de salud en los dominios geográficos donde la variable de la salud no es significativa, como lo son los dominios de la Selva, la Costa,

a excepción de Lima Metropolitana, y la Sierra Centro; de tal forma que, al haber mayores centros de salud, se generará un mayor acceso de los ciudadanos a establecimientos de salud y un potencial efecto marginal sobre la productividad.

- Destacamos la relevancia de la descentralización del gasto público en salud con el fin establecer una distribución más equitativa de servicios públicos de salud disponibles para los ciudadanos de todas las regiones del país. La descentralización, en ese sentido, implica un cambio significativo en la estructura y los flujos de financiamiento; sin embargo, sostenemos que ello es necesario para lograr el acceso, la participación, y la equidad en el sector salud. Los resultados hallados en la presente investigación permiten verificar que no hay una relevancia significativa de la salud en los quintiles más pobres, ello representa un aspecto de mejora para los diversos actores públicos, quienes deberían promover la transferencia de fondos públicos hacia las regiones más pobres y vulnerables del país, donde los beneficios de estas políticas se reflejarán en una mejora de la productividad en el largo plazo.
- Finalmente, recomendamos que las instituciones, tanto privadas como públicas, implementen estrategias eficientes para la prevención de riesgos laborales que atenten contra la salud de los colaboradores; así como, estrategias que mejoren el clima organizacional. Ello mejorará la calidad de vida de los trabajadores, quienes gozarán de un buen estado de salud, y ello se reflejará en menores índices de ausentismo, lo cual generará menores pérdidas económicas para los empleadores.

REFERENCIAS

- Alcázar, L., Ocampo, D., Huamán-Espino, L., y Aparco, J. (2013). Impacto económico de la desnutrición crónica, aguda y global en el Perú. *Revista Peruana de Medicina Experimental y Salud Pública*, 30, 569-574.
- Becker, G. (1964). *Human Capital* (1.^a ed.). Nueva York, Estados Unidos: Columbia University Press for the National Bureau of Economic Research.
- Behrman, J., y Deolalikar, A. (1988). Health and nutrition. *Handbook of development economics*, 1, 631-711.
- Castiñeira, B., Nunes, L., y Rungo, P. (2008). La transmisión de las desigualdades en el estado de salud: efectos sobre la equidad intergeneracional a partir de la acumulación de capital humano. *Cuadernos económicos de ICE*, (75), 99-114.
- Chakraborty, S., y Das, M. (febrero, 2005). Mortality, human capital and persistent inequality. *Journal of Economic growth*, 10(2), 159-192. doi: 10.1007/s10887-005-1670-5.
- Cortez, R. (1999). *Salud y productividad en el Perú: un análisis empírico por género y región*. Banco Interamericano de Desarrollo.
- Domínguez, A. C., González, S. M., Ponzón, A. M. G., Pimienta, L. T., Blanco, Y. E., & Mercado-Caruso, N. (2022). Análisis del nivel de productividad por ausentismo en las empresas. Revisión de la literatura. *Boletín de Innovación, Logística y Operaciones*, 4(1).
- Ehrlich, I., y Lui, F. (octubre, 1991). Intergenerational trade, longevity, and economic growth. *Journal of Political Economy*, 99(5), 1029-1059. doi: 10.1086/261788.
- Fogel, R. (febrero, 1994). Economic growth, population theory, and physiology: the bearing of long-term processes on the making of economic policy. *The American Economic Review*, 84(3), 369-395. doi: 10.3386/w4638.
- González, E., y Hernández, F. (2016). La pérdida de productividad laboral atribuible al tabaquismo. *Revista Cubana de Salud y Trabajo*, 17(2), 57-60.

- Gupta, M., y Barman, T. (junio, 2010). Health, infrastructure, environment and endogenous growth. *Journal of Macroeconomics*, 32(2), 657-673. doi: 10.1016/j.jmacro.2009.09.001.
- Grossman, M. (marzo – abril, 1972). On the concept of health capital and the demand for health. *Journal of Political Economy*, 80(2), 223-255. doi: 10.1086/259880.
- Heckman, J. J. (1979). Sample selection bias as a specification error. *Econometrica: Journal of the econometric society*, 153-161.
- Instituto Nacional de Estadística e Informática (INEI). Base de datos: ENAHO (2019).
- Knaul, F. (1999). *Linking health, nutrition and wages: The evolution of age at menarche and labor earnings among adult Mexican women*. Washington D. C., Estados Unidos: Pan American Health Organization.
- Lewis, A. (septiembre, 1980). The slowing down of the engine of growth. *The American Economic Review*, 70(4), 555-564.
- Mincer, J. (1974). *Schooling, Experience and Earnings*. New York: NBER
- Murrugarra, E., y Valdivia, M. (1999). The returns to health for peruvian urban adults: differentials across genders, the life cycle and the wage distribution. Banco Interamericano de Desarrollo.
- Mushkin, S. (octubre, 1962). Investment in human beings. *Journal of Political Economy*, 70(5), 129-157.
- Monterubbianesi, P. (2014). El Rol de la Salud en el Proceso de Crecimiento Económico: Una Revisión de la Literatura. *Cuadernos de Economía*, 33(62), 91-121. Doi: 10.15446/cuad.econ.v33n62.43667.
- Parker, S. (1999). Elderly health and salaries in the Mexican labor market. Washington D. C., Estados Unidos: InterAmerican Development Bank.
- Pitt, M., Rosenzweig, M., y Hassan, M. (diciembre, 1990). Productivity, health and inequality in the intrahousehold distribution of food in low-income countries. *American Economic Review* 80(5), 1139-1156.

- Ribero, R., y Nuñez, J. (1999). Productivity of household investment in Health: The case of Colombia. Washington D. C., Estados Unidos: InterAmerican Development Bank.
- Ribero, R. (marzo, 2000). Salud y productividad laboral en Colombia. *Revista Desarrollo y sociedad*, (45), 1-30. doi: 10.13043/dys.45.1.
- Sahn, D., y Alderman, H. (1988). The Effects of Variables de Human Capital on Wages, and the Determinants of Labor Supply in a Developing Country. *Journal of Development Economics*, 29(2), 157-183.
- Schultz, T. (marzo, 1961). Investment in Human Capital. *The American Economic Review*, 51(1), 1-17.
- Schultz, T. (1983). Inversión en capital humano. *Educación y sociedad*, 8(3), 180-195.
- Schultz, T., y Tansel, A. (agosto, 1997). Wage and labor supply effects of illness in Cote d'Ivoire and Ghana: instrumental variable estimates for days disabled. *Journal of development economics*, 53(2), 251-286. doi: 10.1016/S0304-3878(97)00025-4.
- Smith, Adam (1776). Investigación de la naturaleza y causas de la riqueza de las naciones. Alianza, Madrid.
- Thomas, D., y John S. (marzo, 1997). Health and wages: Evidence on men and women in urban Brazil. *Journal of Econometrics* 77 (1), 159-186. doi: 10.1016/S0304-4076(96)01811-8.

FUENTES DE DATOS

La base de datos tanto para nuestra variable dependiente como para nuestras variables independientes serán recabadas de la Encuesta Nacional de Hogares sobre las Condiciones de Vida y Pobreza (2019), que se encuentra en la base de datos del INEI: <http://inei.inei.gob.pe/microdatos/>

Delimitación temporal: Nuestra investigación tomará como año de estudio al 2019, que es la fecha del último registro que se tiene sobre la Encuesta Nacional de Hogares sobre Condiciones de Vida y Pobreza; por tanto, nuestro análisis sería lo más actualizado posible y con mayor relevancia para el estudio.

Delimitación geográfica: Se consideran 36.994 hogares, correspondiendo 23.346 al área urbana y 13.648 al área rural del Perú, específicamente de los 24 departamentos del país y la provincia constitucional del Callao. En términos individuales, abarca 121.623 personas.



ANEXOS

Anexo 1: Matriz de Consistencia

Problema o pregunta de investigación	Objetivo	Hipótesis	Variables	Indicadores	Métodos	¿Qué fuentes de datos emplearán?
<p>Pregunta general:</p> <p>¿Cuál es el impacto de la salud sobre la productividad en el Perú, para el año 2019?</p>	<p>Objetivo general:</p> <p>Determinar el impacto de la salud sobre la productividad en el Perú para el año 2019.</p>	<p>Hipótesis general:</p> <p>El impacto de la salud sobre la productividad, en el Perú, es positivo y significativo durante el año de estudio.</p>	<p>Variable dependiente</p> <ul style="list-style-type: none"> - Salario <p>Variable independiente</p> <ul style="list-style-type: none"> - Salud <p>Variables de control</p> <ul style="list-style-type: none"> - Edad - Educación - Experiencia - Desempleo - Sexo - Riqueza - Condiciones de vivienda - Tipo de ocupación 	<p>Variable dependiente</p> <ul style="list-style-type: none"> - Ingreso mensual <p>Variable independiente</p> <ul style="list-style-type: none"> - Indicador de salud <p>Variabes de control</p> <ul style="list-style-type: none"> - Edad - Cuadrado de la edad - Número de años de educación - Cuadrado de los años de educación - Años de experiencia - Cuadrado de los años de experiencia - Sexo - Horas de oferta de agua - Acceso a red telefónica - Sistema de desagüe adecuado - Conexión a internet - Total horas trabajadas 	<p>Mínimos Cuadrados Ordinarios (MCO) corregido por el procedimiento de estimación de Heckman en dos etapas.</p>	<p>Encuesta Nacional de Hogares sobre las Condiciones de Vida y Pobreza (2019).</p>

Problema o pregunta de investigación	Objetivo	Hipótesis	Variables	Indicadores	Métodos	¿Qué fuentes de datos emplearán?
<p>Pregunta específica 1: ¿En qué medida se diferencia el impacto de la salud sobre la productividad entre los ocho dominios geográficos: Costa Norte, Costa Centro, Costa Sur, Sierra Norte, Sierra Centro, Sierra Sur, Selva y Lima Metropolitana?</p>	<p>Objetivo específico 1: Determinar la diferencia del impacto de la salud sobre la productividad entre los ocho dominios geográficos: Costa Norte, Costa Centro, Costa Sur, Sierra Norte, Sierra Centro, Sierra Sur, Selva y Lima Metropolitana.</p>	<p>Hipótesis específica 1: La diferencia del impacto de la salud sobre la productividad, entre los ocho dominios geográficos, es mayor para los dominios situados en la zona sierra y selva, en comparación con la zona costa y Lima Metropolitana la zona limeña.</p>	<p>Variables</p> <ul style="list-style-type: none"> - Salario - Edad - Educación - Experiencia - Desempleo - Sexo - Riqueza - Condiciones de vivienda - Tipo de ocupación 	<p>Variables endógenas</p> <ul style="list-style-type: none"> - Ingreso mensual - Edad - Cuadrado de la edad - Número de años de educación - Cuadrado de los años de educación - Años de experiencia - Cuadrado de los años de experiencia - Sexo - Horas de oferta de agua - Acceso a red telefónica - Sistema de desagüe adecuado - Conexión a internet - Total horas trabajadas 	<p>Mínimos Cuadrados Ordinarios (MCO) corregido por el procedimiento de estimación de Heckman en dos etapas.</p>	<p>Encuesta Nacional de Hogares sobre las Condiciones de Vida y Pobreza (2019).</p>
<p>Pregunta específica 2: ¿En qué medida se diferencia el impacto de la salud sobre la productividad entre los cinco rangos de edad: 18-24, 25-34, 35-44, 45-59, 60-70?</p>	<p>Objetivo específico 2: Determinar la diferencia del impacto de la salud sobre la productividad entre los cinco rangos de edad: 18-24, 25-34, 35-44, 45-59, 60-70.</p>	<p>Hipótesis específica 2: La diferencia del impacto de la salud sobre la productividad, entre los cinco rangos de edad, es mayor para los rangos de edad más jóvenes ya que generalmente se dedican a labores que demandan mayor esfuerzo físico, en comparación con los más adultos que generalmente se dedican a labores que demandan, principalmente, más experiencia.</p>	<p>Variables</p> <ul style="list-style-type: none"> - Salario - Edad - Educación - Experiencia - Desempleo - Sexo - Riqueza - Condiciones de vivienda - Tipo de ocupación 	<p>Variables endógenas</p> <ul style="list-style-type: none"> - Ingreso mensual - Edad - Cuadrado de la edad - Número de años de educación - Cuadrado de los años de educación - Años de experiencia - Cuadrado de los años de experiencia - Sexo - Horas de oferta de agua - Acceso a red telefónica - Sistema de desagüe adecuado - Conexión a internet - Total horas trabajadas 	<p>Mínimos Cuadrados Ordinarios (MCO) corregido por el procedimiento de estimación de Heckman en dos etapas.</p>	<p>Encuesta Nacional de Hogares sobre las Condiciones de Vida y Pobreza (2019).</p>

Problema o pregunta de investigación	Objetivo	Hipótesis	Variables	Indicadores	Métodos	¿Qué fuentes de datos emplearán?
<p>Pregunta específica 3: ¿En qué medida se diferencia el impacto de la salud sobre la productividad entre los quintiles de ingreso?</p>	<p>Objetivo específico 3: Determinar la diferencia del impacto de la salud sobre la productividad entre los quintiles de ingreso.</p>	<p>Hipótesis específica 3: La diferencia del impacto de la salud sobre la productividad, entre los quintiles de ingreso, es mayor para los quintiles más pobres ya que cuentan con un menor presupuesto para invertir en su salud, por tanto, están más propensos a estar enfermos y padecer más dolencias que afecten su productividad.</p>	<p>Variables</p> <ul style="list-style-type: none"> - Salario - Edad - Educación - Experiencia - Desempleo - Sexo - Riqueza - Condiciones de vivienda - Tipo de ocupación 	<p>Variables endógenas</p> <ul style="list-style-type: none"> - Ingreso mensual - Edad - Cuadrado de la edad - Número de años de educación - Cuadrado de los años de educación - Años de experiencia - Cuadrado de los años de experiencia - Sexo - Horas de oferta de agua - Acceso a red telefónica - Sistema de desagüe adecuado - Conexión a internet - Total horas trabajadas 	<p>Mínimos Cuadrados Ordinarios (MCO) corregido por el procedimiento de estimación de Heckman en dos etapas.</p>	<p>Encuesta Nacional de Hogares sobre las Condiciones de Vida y Pobreza (2019).</p>
<p>Pregunta específica 4: ¿En qué medida se diferencia el impacto de la salud sobre la productividad entre los tres niveles educativos: primaria, secundaria, superior?</p>	<p>Objetivo específico 4: Determinar la diferencia del impacto de la salud sobre la productividad entre los tres niveles educativos: primaria, secundaria y superior.</p>	<p>Hipótesis específica 4: La diferencia del impacto de la salud sobre la productividad, entre los tres niveles educativos, es mayor para los niveles educativos más básicos, en ese sentido, habrá un mayor impacto en el nivel primaria, en comparación con el nivel secundaria y superior; y, habrá un mayor impacto en el nivel secundaria, en comparación con el nivel superior.</p>	<p>Variables</p> <ul style="list-style-type: none"> - Salario - Edad - Educación - Experiencia - Desempleo - Sexo - Riqueza - Condiciones de vivienda - Tipo de ocupación 	<p>Variables endógenas</p> <ul style="list-style-type: none"> - Ingreso mensual - Edad - Cuadrado de la edad - Número de años de educación - Cuadrado de los años de educación - Años de experiencia - Cuadrado de los años de experiencia - Sexo - Horas de oferta de agua - Acceso a red telefónica - Sistema de desagüe adecuado - Conexión a internet - Total horas trabajadas 	<p>Mínimos Cuadrados Ordinarios (MCO) corregido por el procedimiento de estimación de Heckman en dos etapas.</p>	<p>Encuesta Nacional de Hogares sobre las Condiciones de Vida y Pobreza (2019).</p>

Anexo 2: Regresión General mediante el modelo de Heckman

Linear regression		Number of obs	=	22,061
		F(14, 22046)	=	681.07
		Prob > F	=	0.0000
		R-squared	=	0.3973
		Root MSE	=	909.95

IM	Robust		t	P> t	[95% Conf. Interval]	
	Coef.	Std. Err.				
E	30.52055	4.116875	7.41	0.000	22.45118	38.58992
E2	-.2910548	.0506269	-5.75	0.000	-.3902873	-.1918224
ED	-134.6329	5.557824	-24.22	0.000	-145.5266	-123.7392
ED2	12.84798	.3156217	40.71	0.000	12.22934	13.46662
EX	25.87606	2.465928	10.49	0.000	21.04266	30.70945
EX2	-.1756936	.0817084	-2.15	0.032	-.3358479	-.0155393
IDS	29.64546	13.36378	2.22	0.027	3.45149	55.83943
THT	8.191809	.3482279	23.52	0.000	7.509258	8.874361
S	243.2421	12.62272	19.27	0.000	218.5006	267.9835
BD	-34.1372	11.90965	-2.87	0.004	-57.48097	-10.79343
OA	6.454253	.5579634	11.57	0.000	5.360604	7.547901
TF	197.3815	22.70772	8.69	0.000	152.8728	241.8903
CI	202.225	13.07992	15.46	0.000	176.5874	227.8626
imr	-2166.518	196.2584	-11.04	0.000	-2551.199	-1781.838
_cons	-244.0871	132.7042	-1.84	0.066	-504.1969	16.02269

Fuente: Elaboración propia

Anexo 3: Test de White para la regresión general

White's test for H_0 : homoskedasticity
against H_a : unrestricted heteroskedasticity

chi2(111) = 1304.51
Prob > chi2 = 0.0000

Cameron & Trivedi's decomposition of IM-test

Source	chi2	df	p
Heteroskedasticity	1304.51	111	0.0000
Skewness	221.31	14	0.0000
Kurtosis	16.03	1	0.0001
Total	1541.85	126	0.0000

Fuente: Elaboración propia

Anexo 4: Prueba de Multicolinealidad para la regresión general

Variable	VIF	1/VIF
E	107.04	0.009342
E2	101.57	0.009846
ED	16.79	0.059567
ED2	15.31	0.065304
EX	9.67	0.103361
EX2	9.26	0.108007
imr	4.07	0.245422
CI	1.53	0.655479
TF	1.38	0.723600
BD	1.31	0.760931
OA	1.11	0.904338
THT	1.05	0.950589
S	1.05	0.952489
IDS	1.00	0.996600
Mean VIF	19.44	

Fuente: Elaboración propia

Anexo 5: Regresión sobre el rango de edad 18-24 mediante el modelo de Heckman

```
. regress IM E E2 ED ED2 EX EX2 IDS THT S BD OA TF CI imr, r
```

```
Linear regression               Number of obs   =       4,121
                               F(14, 4106)     =       83.60
                               Prob > F              =       0.0000
                               R-squared             =       0.2558
                               Root MSE         =       482.94
```

IM	Robust		t	P> t	[95% Conf. Interval]	
	Coef.	Std. Err.				
E	-219.4087	94.71729	-2.32	0.021	-405.1059	-33.71146
E2	4.57622	2.209875	2.07	0.038	.2436669	8.908773
ED	-55.58735	17.98608	-3.09	0.002	-90.84982	-20.32488
ED2	5.31953	.9162317	5.81	0.000	3.523219	7.115841
EX	30.36331	10.81575	2.81	0.005	9.15858	51.56804
EX2	-5.472175	1.344116	-4.07	0.000	-8.107371	-2.836979
IDS	23.68164	16.92922	1.40	0.162	-9.508807	56.87209
THT	9.05469	.4607442	19.65	0.000	8.151382	9.957998
S	64.5755	15.22509	4.24	0.000	34.72608	94.42492
BD	-34.67871	16.80232	-2.06	0.039	-67.62037	-1.737051
OA	4.392884	.7197015	6.10	0.000	2.981879	5.803889
TF	60.17791	27.76272	2.17	0.030	5.747926	114.6079
CI	99.66038	18.82246	5.29	0.000	62.75816	136.5626
imr	-1673.447	341.9503	-4.89	0.000	-2343.855	-1003.039
_cons	3054.989	1085.593	2.81	0.005	926.6374	5183.34

Fuente: Elaboración propia

Anexo 6: Test de White para la regresión sobre el rango de edad 18-24

White's test for H_0 : homoskedasticity
against H_a : unrestricted heteroskedasticity

chi2(111) = 332.78
Prob > chi2 = 0.0000

Cameron & Trivedi's decomposition of IM-test

Source	chi2	df	p
Heteroskedasticity	332.78	111	0.0000
Skewness	85.25	14	0.0000
Kurtosis	13.57	1	0.0002
Total	431.59	126	0.0000

Fuente: Elaboración propia

Anexo 7: Prueba de Multicolinealidad para la regresión sobre el rango de edad 18-24

Variable	VIF	1/VIF
E	656.30	0.001524
E2	610.60	0.001638
ED2	31.00	0.032262
ED	30.78	0.032490
imr	17.76	0.056297
EX	4.69	0.213193
EX2	4.67	0.214188
CI	1.40	0.712461
TF	1.29	0.774386
BD	1.28	0.778824
OA	1.05	0.951325
S	1.04	0.957947
THT	1.04	0.964146
IDS	1.01	0.989666
Mean VIF	97.42	

Fuente: Elaboración propia

Anexo 8: Regresión sobre el rango de edad 25-34 mediante el modelo de Heckman

Linear regression	Number of obs	=	5,708
	F(14, 5693)	=	149.04
	Prob > F	=	0.0000
	R-squared	=	0.3389
	Root MSE	=	876.23

IM	Robust		t	P> t	[95% Conf. Interval]	
	Coef.	Std. Err.				
E	226.8583	96.24244	2.36	0.018	38.18644	415.5301
E2	-3.209683	1.645576	-1.95	0.051	-6.435639	.0162719
ED	-243.8657	16.15695	-15.09	0.000	-275.5394	-212.1919
ED2	16.79306	.8124409	20.67	0.000	15.20037	18.38576
EX	75.64647	8.724791	8.67	0.000	58.54256	92.75039
EX2	-5.028167	.5972328	-8.42	0.000	-6.198971	-3.857363
IDS	36.91408	26.22201	1.41	0.159	-14.49104	88.31921
THT	8.383016	.6987449	12.00	0.000	7.013209	9.752822
S	227.561	23.93182	9.51	0.000	180.6456	274.4765
BD	40.52896	22.37957	1.81	0.070	-3.343523	84.40144
OA	6.210525	1.08047	5.75	0.000	4.092393	8.328658
TF	107.1541	42.56703	2.52	0.012	23.70655	190.6017
CI	227.5854	25.97524	8.76	0.000	176.664	278.5068
imr	-949.1636	245.1769	-3.87	0.000	-1429.804	-468.5236
_cons	-3147.686	1410.436	-2.23	0.026	-5912.678	-382.6943

Fuente: Elaboración propia

Anexo 9: Test de White para la regresión sobre el rango de edad 25-34

White's test for H₀: homoskedasticity
against H_a: unrestricted heteroskedasticity

chi2(111) = 303.55
Prob > chi2 = 0.0000

Cameron & Trivedi's decomposition of IM-test

Source	chi2	df	p
Heteroskedasticity	303.55	111	0.0000
Skewness	55.91	14	0.0000
Kurtosis	5.33	1	0.0210
Total	364.79	126	0.0000

Fuente: Elaboración propia

Anexo 10: Prueba de Multicolinealidad para la regresión sobre el rango de edad 25-34

Variable	VIF	1/VIF
E	571.68	0.001749
E2	571.17	0.001751
ED	27.75	0.036036
ED2	26.97	0.037083
EX2	5.53	0.180942
EX	5.46	0.183082
imr	1.88	0.533214
CI	1.51	0.664073
TF	1.34	0.746338
BD	1.30	0.767048
OA	1.07	0.933473
S	1.05	0.953344
THT	1.05	0.956647
IDS	1.01	0.994313
Mean VIF	87.05	

Fuente: Elaboración propia

Anexo 11: Regresión sobre el rango de edad 35-44 mediante el modelo de Heckman

Linear regression	Number of obs	=	5,102
	F(14, 5087)	=	177.52
	Prob > F	=	0.0000
	R-squared	=	0.4029
	Root MSE	=	972.57

IM	Robust		t	P> t	[95% Conf. Interval]	
	Coef.	Std. Err.				
E	255.0548	146.4597	1.74	0.082	-32.0693	542.1789
E2	-3.061071	1.854741	-1.65	0.099	-6.697161	.5750198
ED	-291.8112	19.57622	-14.91	0.000	-330.189	-253.4334
ED2	20.42651	.985727	20.72	0.000	18.49406	22.35896
EX	52.97741	6.081003	8.71	0.000	41.05603	64.89879
EX2	-2.082265	.2750217	-7.57	0.000	-2.621425	-1.543104
IDS	64.64259	29.15872	2.22	0.027	7.478938	121.8062
THT	7.249334	.8315329	8.72	0.000	5.619171	8.879496
S	382.2687	28.94168	13.21	0.000	325.5305	439.0068
BD	9.002586	24.93678	0.36	0.718	-39.88424	57.88942
OA	7.238315	1.239788	5.84	0.000	4.807796	9.668833
TF	240.3119	54.89238	4.38	0.000	132.6992	347.9246
CI	307.9127	30.64514	10.05	0.000	247.835	367.9903
imr	-1697.566	378.4478	-4.49	0.000	-2439.486	-955.6452
_cons	-4317.523	2874.608	-1.50	0.133	-9952.992	1317.946

Fuente: Elaboración propia

Anexo 12: Test de White para la regresión sobre el rango de edad 35-44

White's test for Ho: homoskedasticity
 against Ha: unrestricted heteroskedasticity

chi2(111) = 378.42
 Prob > chi2 = 0.0000

Cameron & Trivedi's decomposition of IM-test

Source	chi2	df	p
Heteroskedasticity	378.42	111	0.0000
Skewness	81.50	14	0.0000
Kurtosis	9.41	1	0.0022
Total	469.33	126	0.0000

Fuente: Elaboración propia

Anexo 13: Prueba de Multicolinealidad para la regresión sobre el rango de edad 35-44

Variable	VIF	1/VIF
E2	984.44	0.001016
E	984.29	0.001016
ED	38.00	0.026318
ED2	33.68	0.029689
EX	7.64	0.130948
EX2	7.63	0.131115
imr	2.60	0.385206
CI	1.48	0.673881
TF	1.32	0.756633
BD	1.30	0.771483
OA	1.14	0.878093
S	1.08	0.925547
THT	1.08	0.927319
IDS	1.00	0.998156
Mean VIF	147.62	

Fuente: Elaboración propia

Anexo 14: Regresión sobre el rango de edad 45-59 mediante el modelo de Heckman

Linear regression		Number of obs	=	5,503
		F(14, 5488)	=	267.94
		Prob > F	=	0.0000
		R-squared	=	0.4295
		Root MSE	=	982.33

IM	Robust		t	P> t	[95% Conf. Interval]	
	Coef.	Std. Err.				
E	5.902742	81.7756	0.07	0.942	-154.4099	166.2153
E2	-.1611497	.789586	-0.20	0.838	-1.709051	1.386752
ED	-171.4646	10.62526	-16.14	0.000	-192.2943	-150.6349
ED2	15.93096	.6624828	24.05	0.000	14.63223	17.22968
EX	32.30052	4.058381	7.96	0.000	24.34449	40.25656
EX2	-.504704	.1260538	-4.00	0.000	-.7518194	-.2575886
IDS	-4.435448	28.29514	-0.16	0.875	-59.90514	51.03424
THT	6.661319	.7503642	8.88	0.000	5.190308	8.13233
S	276.893	27.37642	10.11	0.000	223.2244	330.5616
BD	-38.0631	27.34022	-1.39	0.164	-91.66077	15.53457
OA	4.80878	1.233599	3.90	0.000	2.390437	7.227124
TF	257.7342	42.9473	6.00	0.000	173.5405	341.9279
CI	208.1215	26.5795	7.83	0.000	156.0152	260.2279
imr	-2717.022	365.6714	-7.43	0.000	-3433.883	-2000.162
_cons	864.4001	2111.29	0.41	0.682	-3274.566	5003.366

Fuente: Elaboración propia

Anexo 15: Test de White para la regresión sobre el rango de edad 45-59

White's test for H₀: homoskedasticity
against H_a: unrestricted heteroskedasticity

chi2(111) = 434.78
Prob > chi2 = 0.0000

Cameron & Trivedi's decomposition of IM-test

Source	chi2	df	p
Heteroskedasticity	434.78	111	0.0000
Skewness	89.08	14	0.0000
Kurtosis	7.94	1	0.0048
Total	531.80	126	0.0000

Fuente: Elaboración propia

Anexo 16: Prueba de Multicolinealidad para la regresión sobre el rango de edad 45-59

Variable	VIF	1/VIF
E2	722.98	0.001383
E	722.64	0.001384
ED	20.30	0.049259
ED2	20.20	0.049516
EX	12.04	0.083090
EX2	11.88	0.084209
imr	2.23	0.448883
CI	1.58	0.631403
TF	1.43	0.700797
BD	1.30	0.771166
OA	1.26	0.795043
THT	1.09	0.914122
S	1.07	0.936251
IDS	1.01	0.993219
Mean VIF	108.64	

Fuente: Elaboración propia

Anexo 17: Regresión sobre el rango de edad 60-70 mediante el modelo de Heckman

Linear regression	Number of obs	=	1,627
	F(14, 1612)	=	78.73
	Prob > F	=	0.0000
	R-squared	=	0.4187
	Root MSE	=	1125.9

IM	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
E	117.2769	425.6692	0.28	0.783	-717.6463	952.2002
E2	-.9036303	3.295045	-0.27	0.784	-7.366652	5.559391
ED	-167.3005	20.53352	-8.15	0.000	-207.5757	-127.0253
ED2	16.37415	1.519048	10.78	0.000	13.39464	19.35367
EX	9.376755	7.392688	1.27	0.205	-5.123534	23.87704
EX2	.3283785	.1802041	1.82	0.069	-.0250804	.6818373
IDS	11.81447	63.06777	0.19	0.851	-111.889	135.5179
THT	10.18007	1.412865	7.21	0.000	7.408827	12.95132
S	305.7561	61.76415	4.95	0.000	184.6096	426.9026
BD	108.2306	47.86408	2.26	0.024	14.34824	202.113
OA	8.609264	2.487304	3.46	0.001	3.730575	13.48795
TF	231.3875	82.33398	2.81	0.005	69.89464	392.8804
CI	254.9844	54.25377	4.70	0.000	148.569	361.3997
imr	-1886.474	365.878	-5.16	0.000	-2604.121	-1168.828
_cons	-3776.567	13752.04	-0.27	0.784	-30750.32	23197.19

Fuente: Elaboración propia

Anexo 18: Test de White para la regresión sobre el rango de edad 60-70

White's test for H_0 : homoskedasticity
against H_a : unrestricted heteroskedasticity

chi2(111) = 130.76
Prob > chi2 = 0.0970

Cameron & Trivedi's decomposition of IM-test

Source	chi2	df	p
Heteroskedasticity	130.76	111	0.0970
Skewness	22.10	14	0.0765
Kurtosis	2.03	1	0.1543
Total	154.89	126	0.0411

Fuente: Elaboración propia

Anexo 19: Prueba de Multicolinealidad para la regresión sobre el rango de edad 60-70

Variable	VIF	1/VIF
E2	2245.72	0.000445
E	2240.73	0.000446
ED	19.93	0.050178
ED2	19.64	0.050915
EX	14.18	0.070545
EX2	13.79	0.072509
imr	1.73	0.576930
CI	1.60	0.624994
TF	1.50	0.667615
BD	1.23	0.815225
OA	1.15	0.867376
S	1.08	0.922647
THT	1.08	0.926837
IDS	1.01	0.989103
Mean VIF	326.03	

Fuente: Elaboración propia

Anexo 20: Regresión sobre el primer quintil mediante el modelo de Heckman

Linear regression

Number of obs	=	4,412
F(14, 4397)	=	74.57
Prob > F	=	0.0000
R-squared	=	0.2387
Root MSE	=	50.581

IM	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
E	.3909998	.6132048	0.64	0.524	-.8111904	1.59319
E2	-.004558	.0073009	-0.62	0.532	-.0188715	.0097555
ED	-.6376164	.9365827	-0.68	0.496	-2.47379	1.198557
ED2	.1068099	.0713505	1.50	0.134	-.0330729	.2466927
EX	.5852683	.2723235	2.15	0.032	.0513771	1.119159
EX2	-.0085467	.0084633	-1.01	0.313	-.0251391	.0080458
IDS	.8916375	1.770397	0.50	0.615	-2.579232	4.362507
THT	1.405361	.0485688	28.94	0.000	1.310142	1.50058
S	17.28465	1.60471	10.77	0.000	14.13861	20.43069
BD	-.2062385	1.716157	-0.12	0.904	-3.57077	3.158293
OA	-.0103068	.0757636	-0.14	0.892	-.1588415	.138228
TF	2.544792	3.411309	0.75	0.456	-4.143092	9.232676
CI	7.120415	2.10502	3.38	0.001	2.993516	11.24731
imr	-49.19908	26.44043	-1.86	0.063	-101.0356	2.637468
_cons	76.97391	24.06953	3.20	0.001	29.78551	124.1623

Fuente: Elaboración propia

Anexo 21: Test de White para la regresión sobre el primer quintil

White's test for H_0 : homoskedasticity
against H_a : unrestricted heteroskedasticity

chi2(111) = 326.26
Prob > chi2 = 0.0000

Cameron & Trivedi's decomposition of IM-test

Source	chi2	df	p
Heteroskedasticity	326.26	111	0.0000
Skewness	771.50	14	0.0000
Kurtosis	3.69	1	0.0547
Total	1101.45	126	0.0000

Fuente: Elaboración propia

Anexo 22: Prueba de Multicolinealidad para la regresión sobre el primer quintil

Variable	VIF	1/VIF
E	136.14	0.007346
E2	123.95	0.008068
ED2	36.51	0.027391
ED	26.70	0.037454
imr	8.97	0.111500
EX	7.09	0.141022
EX2	6.89	0.145129
CI	1.29	0.777173
BD	1.24	0.804977
TF	1.20	0.830591
OA	1.11	0.903126
S	1.09	0.917056
THT	1.02	0.975925
IDS	1.01	0.987458
Mean VIF	25.30	

Fuente: Elaboración propia

Anexo 23: Regresión sobre el segundo quintil mediante el modelo de Heckman

Linear regression		Number of obs	=	4,412
		F(14, 4397)	=	15.02
		Prob > F	=	0.0000
		R-squared	=	0.0450
		Root MSE	=	87.818

IM	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]
E	-.5321576	.6787473	-0.78	0.433	-1.862844 .7985289
E2	.0053078	.0083062	0.64	0.523	-.0109766 .0215921
ED	-1.208281	2.500102	-0.48	0.629	-6.109741 3.693178
ED2	.1816735	.1816872	1.00	0.317	-.174525 .5378719
EX	-.7297851	.4976549	-1.47	0.143	-1.705439 .2458691
EX2	.0114739	.0170139	0.67	0.500	-.0218821 .0448298
IDS	6.804718	3.05411	2.23	0.026	.8171237 12.79231
THT	-.3074932	.0860444	-3.57	0.000	-.4761836 -.1388029
S	-18.3587	2.910723	-6.31	0.000	-24.06518 -12.65222
BD	11.09747	3.167901	3.50	0.000	4.886794 17.30816
OA	.3761079	.1491134	2.52	0.012	.0837705 .6684453
TF	8.850649	5.00449	1.77	0.077	-.9606716 18.66197
CI	1.363792	3.273972	0.42	0.677	-5.054842 7.782426
imr	8.799189	78.68594	0.11	0.911	-145.4649 163.0633
_cons	357.5446	25.61162	13.96	0.000	307.333 407.7563

Fuente: Elaboración propia

Anexo 24: Test de White para la regresión sobre el segundo quintil

White's test for H_0 : homoskedasticity
against H_a : unrestricted heteroskedasticity

chi2 (111) = 166.54
Prob > chi2 = 0.0005

Cameron & Trivedi's decomposition of IM-test

Source	chi2	df	p
Heteroskedasticity	166.54	111	0.0005
Skewness	334.17	14	0.0000
Kurtosis	484.06	1	0.0000
Total	984.77	126	0.0000

Fuente: Elaboración propia

Anexo 25: Prueba de Multicolinealidad para la regresión sobre el segundo quintil

Variable	VIF	1/VIF
ED2	77.97	0.012825
ED	55.33	0.018074
E	43.07	0.023216
E2	41.46	0.024120
imr	9.29	0.107623
EX	6.06	0.165128
EX2	5.86	0.170786
OA	1.40	0.713237
CI	1.32	0.755986
TF	1.25	0.799280
BD	1.21	0.823723
THT	1.07	0.932353
S	1.06	0.943834
IDS	1.01	0.993561
Mean VIF	17.67	

Fuente: Elaboración propia

Anexo 26: Regresión sobre el tercer quintil mediante el modelo de Heckman

Linear regression

Number of obs	=	4,412
F(14, 4397)	=	22.38
Prob > F	=	0.0000
R-squared	=	0.0633
Root MSE	=	144.74

IM	Robust		t	P> t	[95% Conf. Interval]	
	Coef.	Std. Err.				
E	-.2071458	1.158143	-0.18	0.858	-2.477689	2.063397
E2	-.004412	.0134602	-0.33	0.743	-.0308007	.0219768
ED	-.4407263	2.139234	-0.21	0.837	-4.634702	3.75325
ED2	.1525845	.161373	0.95	0.344	-.1637877	.4689568
EX	2.122621	.9327717	2.28	0.023	.2939187	3.951323
EX2	-.0557343	.0311431	-1.79	0.074	-.1167905	.0053218
IDS	-2.529183	4.852056	-0.52	0.602	-12.04166	6.98329
THT	1.209784	.1462217	8.27	0.000	.9231155	1.496452
S	4.86915	4.454816	1.09	0.274	-3.864532	13.60283
BD	15.45732	6.103097	2.53	0.011	3.492178	27.42247
OA	1.169557	.2814001	4.16	0.000	.6178715	1.721243
TF	19.39875	6.115485	3.17	0.002	7.409315	31.38818
CI	7.389251	5.10645	1.45	0.148	-2.621963	17.40047
imr	568.4342	233.7983	2.43	0.015	110.0717	1026.797
_cons	655.8485	24.16815	27.14	0.000	608.4668	703.2303

Fuente: Elaboración propia

Anexo 27: Test de White para la regresión sobre el tercer quintil

White's test for H_0 : homoskedasticity
 against H_a : unrestricted heteroskedasticity

chi2(111)	=	196.58
Prob > chi2	=	0.0000

Cameron & Trivedi's decomposition of IM-test

Source	chi2	df	p
Heteroskedasticity	196.58	111	0.0000
Skewness	599.81	14	0.0000
Kurtosis	387.30	1	0.0000
Total	1183.69	126	0.0000

Fuente: Elaboración propia

Anexo 28: Prueba de Multicolinealidad para la regresión sobre el tercer quintil

Variable	VIF	1/VIF
E	47.19	0.021189
E2	42.41	0.023577
ED2	26.01	0.038445
ED	13.07	0.076512
imr	12.08	0.082795
EX	7.27	0.137640
EX2	6.80	0.147119
OA	1.71	0.583859
CI	1.33	0.749541
TF	1.33	0.754277
BD	1.19	0.840216
S	1.05	0.951617
THT	1.05	0.952351
IDS	1.01	0.993045
Mean VIF	11.68	

Fuente: Elaboración propia

Anexo 29: Regresión sobre el cuarto quintil mediante el modelo de Heckman

Linear regression		Number of obs	=	4,412
		F(14, 4397)	=	23.89
		Prob > F	=	0.0000
		R-squared	=	0.0679
		Root MSE	=	211.7

IM	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]
E	7.346586	1.768321	4.15	0.000	3.879786 10.81339
E2	-.0793828	.0205933	-3.85	0.000	-.1197561 -.0390095
ED	-9.53711	3.431942	-2.78	0.005	-16.26544 -2.808776
ED2	1.201546	.1743452	6.89	0.000	.8597415 1.54335
EX	2.836476	1.171017	2.42	0.015	.5406924 5.13226
EX2	-.0097459	.0359331	-0.27	0.786	-.0801928 .060701
IDS	14.96809	7.470363	2.00	0.045	.3224126 29.61376
THT	.2149923	.2118648	1.01	0.310	-.2003695 .6303541
S	67.67629	6.835293	9.90	0.000	54.27567 81.07691
BD	-4.984389	9.445179	-0.53	0.598	-23.5017 13.53292
OA	.572061	.3571418	1.60	0.109	-.1281168 1.272239
TF	8.639601	8.226758	1.05	0.294	-7.488988 24.76819
CI	-12.0925	7.388997	-1.64	0.102	-26.57865 2.393659
imr	-174.6625	221.0365	-0.79	0.429	-608.0054 258.6805
_cons	1076.896	39.2646	27.43	0.000	999.9178 1153.875

Fuente: Elaboración propia

Anexo 30: Test de White para la regresión sobre el cuarto quintil

White's test for H_0 : homoskedasticity
against H_a : unrestricted heteroskedasticity

chi2(111) = 136.25
Prob > chi2 = 0.0521

Cameron & Trivedi's decomposition of IM-test

Source	chi2	df	p
Heteroskedasticity	136.25	111	0.0521
Skewness	275.55	14	0.0000
Kurtosis	260.92	1	0.0000
Total	672.72	126	0.0000

Fuente: Elaboración propia

Anexo 31: Prueba de Multicolinealidad para la regresión sobre el cuarto quintil

Variable	VIF	1/VIF
E	51.72	0.019335
E2	49.27	0.020294
ED2	15.33	0.065234
ED	14.33	0.069791
EX	9.34	0.107050
EX2	8.70	0.114965
imr	4.44	0.225424
CI	1.33	0.749520
TF	1.30	0.770214
OA	1.26	0.796555
BD	1.16	0.860721
S	1.10	0.912782
THT	1.09	0.915326
IDS	1.01	0.994384
Mean VIF	11.53	

Fuente: Elaboración propia

Anexo 32: Regresión sobre el quinto quintil mediante el modelo de Heckman

Linear regression	Number of obs	=	4,413
	F(14, 4398)	=	32.41
	Prob > F	=	0.0000
	R-squared	=	0.1248
	Root MSE	=	1291.3

IM	Robust		t	P> t	[95% Conf. Interval]	
	Coef.	Std. Err.				
E	85.58533	14.30158	5.98	0.000	57.54702	113.6236
E2	-.675706	.1619494	-4.17	0.000	-.9932085	-.3582036
ED	-223.3578	35.67916	-6.26	0.000	-293.307	-153.4087
ED2	14.78011	1.586942	9.31	0.000	11.6689	17.89131
EX	2.94054	6.202115	0.47	0.635	-9.218728	15.09981
EX2	-.1762784	.1668674	-1.06	0.291	-.5034225	.1508656
IDS	107.1451	40.58257	2.64	0.008	27.5828	186.7073
THT	9.023566	1.152706	7.83	0.000	6.763683	11.28345
S	336.118	44.75455	7.51	0.000	248.3765	423.8594
BD	-42.86468	43.88723	-0.98	0.329	-128.9057	43.17639
OA	9.445262	1.786457	5.29	0.000	5.942907	12.94762
TF	287.6102	47.43514	6.06	0.000	194.6135	380.607
CI	207.9197	35.52542	5.85	0.000	138.2719	277.5674
imr	-6455.913	631.0618	-10.23	0.000	-7693.112	-5218.714
_cons	711.9821	334.6052	2.13	0.033	55.98737	1367.977

Fuente: Elaboración propia

Anexo 33: Test de White para la regresión sobre el quinto quintil

White's test for H_0 : homoskedasticity
against H_a : unrestricted heteroskedasticity

chi2(111) = 248.10
Prob > chi2 = 0.0000

Cameron & Trivedi's decomposition of IM-test

Source	chi2	df	p
Heteroskedasticity	248.10	111	0.0000
Skewness	72.91	14	0.0000
Kurtosis	8.81	1	0.0030
Total	329.82	126	0.0000

Fuente: Elaboración propia

Anexo 34: Prueba de Multicolinealidad para la regresión sobre el quinto quintil

Variable	VIF	1/VIF
E	84.65	0.011813
E2	78.13	0.012799
ED2	23.99	0.041684
ED	21.31	0.046930
EX	14.42	0.069331
EX2	14.09	0.070967
imr	4.02	0.248683
OA	1.34	0.745316
CI	1.29	0.773491
TF	1.28	0.782700
THT	1.19	0.841605
S	1.17	0.855711
BD	1.15	0.870519
IDS	1.00	0.995092
Mean VIF	17.79	

Fuente: Elaboración propia

Anexo 35: Regresión sobre el nivel educativo primaria mediante el modelo de Heckman

Linear regression		Number of obs	=	3,142
		F(14, 3127)	=	36.57
		Prob > F	=	0.0000
		R-squared	=	0.1726
		Root MSE	=	462.45

IM	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]
E	5.0278	4.909575	1.02	0.306	-4.598517 14.65412
E2	-.0247003	.0607776	-0.41	0.684	-.1438683 .0944677
ED	16.72965	21.01554	0.80	0.426	-24.47599 57.93529
ED2	-1.917386	2.916339	-0.66	0.511	-7.635518 3.800745
EX	-2.865032	3.105009	-0.92	0.356	-8.953094 3.22303
EX2	.2291855	.1231644	1.86	0.063	-.0123058 .4706767
IDS	-23.28852	17.74254	-1.31	0.189	-58.07674 11.49969
THT	8.740813	.5127284	17.05	0.000	7.735495 9.746132
S	125.9227	15.58778	8.08	0.000	95.35933 156.486
BD	47.6937	16.24185	2.94	0.003	15.84793 79.53948
OA	2.135997	.8513972	2.51	0.012	.4666424 3.805351
TF	31.60063	57.4288	0.55	0.582	-81.00132 144.2026
CI	124.1121	30.30986	4.09	0.000	64.68281 183.5413
imr	-638.7032	175.3399	-3.64	0.000	-982.4962 -294.9102
_cons	-51.06509	125.132	-0.41	0.683	-296.4142 194.284

Fuente: Elaboración propia

Anexo 36: Test de White para la regresión sobre el nivel educativo primaria

White's test for Ho: homoskedasticity
against Ha: unrestricted heteroskedasticity

chi2(111) = 480.15
Prob > chi2 = 0.0000

Cameron & Trivedi's decomposition of IM-test

Source	chi2	df	p
Heteroskedasticity	480.15	111	0.0000
Skewness	90.20	14	0.0000
Kurtosis	3.58	1	0.0586
Total	573.93	126	0.0000

Fuente: Elaboración propia

Anexo 37: Prueba de Multicolinealidad para la regresión sobre el nivel educativo primaria

Variable	VIF	1/VIF
E2	60.34	0.016572
E	51.05	0.019587
ED	39.95	0.025034
ED2	30.98	0.032283
EX2	7.50	0.133273
EX	7.50	0.133364
imr	1.33	0.749144
BD	1.23	0.812805
CI	1.19	0.837986
TF	1.16	0.864017
S	1.10	0.911640
OA	1.09	0.918061
THT	1.06	0.943370
IDS	1.01	0.987460
Mean VIF	14.75	

Fuente: Elaboración propia

Anexo 38: Regresión sobre el nivel educativo secundaria mediante el modelo de Heckman

Linear regression	Number of obs	=	8,822
	F(14, 8807)	=	132.54
	Prob > F	=	0.0000
	R-squared	=	0.2142
	Root MSE	=	655.55

IM	Robust		t	P> t	[95% Conf. Interval]	
	Coef.	Std. Err.				
E	6.681676	4.681095	1.43	0.154	-2.494364	15.85771
E2	-.0420138	.056794	-0.74	0.459	-.1533434	.0693157
ED	80.94029	103.5034	0.78	0.434	-121.9506	283.8312
ED2	-1.17287	5.564847	-0.21	0.833	-12.08127	9.73553
EX	2.253561	3.696376	0.61	0.542	-4.992199	9.499321
EX2	.5731084	.1468462	3.90	0.000	.2852555	.8609612
IDS	3.368181	15.93209	0.21	0.833	-27.86244	34.5988
THT	7.074392	.4273793	16.55	0.000	6.236629	7.912156
S	246.7255	13.64906	18.08	0.000	219.9702	273.4809
BD	45.14135	14.66341	3.08	0.002	16.39764	73.88507
OA	3.90381	.6696928	5.83	0.000	2.591056	5.216564
TF	109.5304	33.18431	3.30	0.001	44.48139	174.5794
CI	184.6845	18.35502	10.06	0.000	148.7043	220.6646
imr	-1591.781	228.885	-6.95	0.000	-2040.449	-1143.113
_cons	-483.5371	448.7193	-1.08	0.281	-1363.132	396.0574

Fuente: Elaboración propia

Anexo 39: Test de White para la regresión sobre el nivel educativo secundaria

White's test for H_0 : homoskedasticity
against H_a : unrestricted heteroskedasticity

chi2(111) = 439.23
Prob > chi2 = 0.0000

Cameron & Trivedi's decomposition of IM-test

Source	chi2	df	p
Heteroskedasticity	439.23	111	0.0000
Skewness	41.38	14	0.0002
Kurtosis	3.02	1	0.0821
Total	483.64	126	0.0000

Fuente: Elaboración propia

Anexo 40: Prueba de Multicolinealidad para la regresión sobre el nivel educativo secundaria

Variable	VIF	1/VIF
ED2	458.19	0.002182
ED	454.88	0.002198
E	86.92	0.011505
E2	79.13	0.012638
EX	7.71	0.129625
EX2	7.42	0.134831
imr	3.26	0.306509
CI	1.30	0.767795
TF	1.28	0.782186
BD	1.21	0.827971
OA	1.07	0.937103
S	1.04	0.957294
THT	1.04	0.965177
IDS	1.00	0.997492
Mean VIF	78.96	

Fuente: Elaboración propia

Anexo 41: Regresión sobre el nivel educativo superior mediante el modelo de Heckman

Linear regression

Number of obs	=	10,097
F(14, 10082)	=	355.41
Prob > F	=	0.0000
R-squared	=	0.3317
Root MSE	=	1131.2

IM	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]
E	27.83092	10.95077	2.54	0.011	6.365233 49.29661
E2	-.2079276	.1390197	-1.50	0.135	-.4804339 .0645786
ED	-3113.897	239.6455	-12.99	0.000	-3583.65 -2644.144
ED2	116.469	8.284043	14.06	0.000	100.2307 132.7074
EX	51.5456	4.708949	10.95	0.000	42.31512 60.77608
EX2	-.8021589	.1394609	-5.75	0.000	-1.07553 -.5287878
IDS	59.81594	23.58046	2.54	0.011	13.59354 106.0384
THT	10.71099	.6814039	15.72	0.000	9.375304 12.04668
S	280.9806	22.60616	12.43	0.000	236.668 325.2932
BD	-42.4329	26.74019	-1.59	0.113	-94.849 9.983197
OA	7.80554	1.153618	6.77	0.000	5.54422 10.06686
TF	180.9568	29.5797	6.12	0.000	122.9747 238.9389
CI	284.1648	20.74174	13.70	0.000	243.5069 324.8228
imr	-3454.269	450.3054	-7.67	0.000	-4336.958 -2571.581
_cons	20901.98	1934.589	10.80	0.000	17109.8 24694.16

Fuente: Elaboración propia

Anexo 42: Test de White para la regresión sobre el nivel educativo superior

White's test for H_0 : homoskedasticity
against H_a : unrestricted heteroskedasticity

chi2(111) = 556.34
Prob > chi2 = 0.0000

Cameron & Trivedi's decomposition of IM-test

Source	chi2	df	p
Heteroskedasticity	556.34	111	0.0000
Skewness	122.33	14	0.0000
Kurtosis	12.61	1	0.0004
Total	691.28	126	0.0000

Fuente: Elaboración propia

Anexo 43: Prueba de Multicolinealidad para la regresión sobre el nivel educativo superior

Variable	VIF	1/VIF
ED	1200.73	0.000833
ED2	1171.92	0.000853
E	217.07	0.004607
E2	213.53	0.004683
EX	12.90	0.077545
EX2	12.11	0.082577
imr	8.91	0.112250
OA	1.48	0.675585
TF	1.31	0.762399
CI	1.30	0.770308
BD	1.16	0.861031
THT	1.06	0.943525
S	1.03	0.972735
IDS	1.00	0.996118
Mean VIF	203.25	

Fuente: Elaboración propia

Anexo 44: Regresión sobre el dominio Sierra Sur mediante el modelo de Heckman

Linear regression		Number of obs	=	2,375
		F(14, 2360)	=	84.10
		Prob > F	=	0.0000
		R-squared	=	0.4137
		Root MSE	=	982.67

IM	Robust		t	P> t	[95% Conf. Interval]	
	Coef.	Std. Err.				
E	3.156723	12.85245	0.25	0.806	-22.04655	28.35999
E2	.0227329	.1545111	0.15	0.883	-.2802587	.3257244
ED	-150.7738	20.33683	-7.41	0.000	-190.6537	-110.8939
ED2	12.7537	1.064923	11.98	0.000	10.66542	14.84199
EX	26.57745	9.162104	2.90	0.004	8.610838	44.54405
EX2	.2100892	.3379927	0.62	0.534	-.4527042	.8728825
IDS	69.14476	40.45744	1.71	0.088	-10.19105	148.4806
THT	13.04397	1.134719	11.50	0.000	10.81882	15.26912
S	437.3861	41.07734	10.65	0.000	356.8347	517.9375
BD	15.45838	39.24339	0.39	0.694	-61.49671	92.41348
OA	3.312058	1.893594	1.75	0.080	-.4012217	7.025338
TF	-54.19407	84.5018	-0.64	0.521	-219.8995	111.5114
CI	322.0361	47.74433	6.75	0.000	228.4109	415.6613
imr	-3021.346	478.5349	-6.31	0.000	-3959.739	-2082.954
_cons	459.1309	408.7202	1.12	0.261	-342.3569	1260.619

Fuente: Elaboración propia

Anexo 45: Test de White para la regresión sobre el dominio Sierra Sur

White's test for H_0 : homoskedasticity
against H_a : unrestricted heteroskedasticity

chi2(111) = 485.96
Prob > chi2 = 0.0000

Cameron & Trivedi's decomposition of IM-test

Source	chi2	df	p
Heteroskedasticity	485.96	111	0.0000
Skewness	134.25	14	0.0000
Kurtosis	15.03	1	0.0001
Total	635.24	126	0.0000

Fuente: Elaboración propia

Anexo 46: Prueba de Multicolinealidad para la regresión sobre el dominio Sierra Sur

Variable	VIF	1/VIF
E	90.12	0.011096
E2	86.79	0.011522
ED	16.51	0.060560
ED2	16.37	0.061105
EX	10.37	0.096447
EX2	9.95	0.100507
imr	3.20	0.312255
CI	1.52	0.656882
OA	1.38	0.724128
TF	1.34	0.746151
BD	1.30	0.768423
S	1.07	0.934349
THT	1.06	0.943332
IDS	1.02	0.979593
Mean VIF	17.29	

Fuente: Elaboración propia

Anexo 47: Regresión sobre el dominio Sierra Norte mediante el modelo de Heckman

Linear regression	Number of obs	=	956
	F(14, 941)	=	85.87
	Prob > F	=	0.0000
	R-squared	=	0.5392
	Root MSE	=	732.03

IM	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
E	40.24932	12.11565	3.32	0.001	16.47249	64.02615
E2	-.3314955	.1630629	-2.03	0.042	-.6515045	-.0114865
ED	-95.52739	22.34214	-4.28	0.000	-139.3736	-51.6812
ED2	10.30908	1.231021	8.37	0.000	7.893211	12.72494
EX	29.80012	8.792931	3.39	0.001	12.5441	47.05614
EX2	-.5586258	.2626748	-2.13	0.034	-1.074122	-.0431296
IDS	-134.4475	55.39814	-2.43	0.015	-243.1657	-25.72931
THT	8.142737	1.348112	6.04	0.000	5.497082	10.78839
S	137.6999	48.75657	2.82	0.005	42.01572	233.3841
BD	-12.70765	51.94746	-0.24	0.807	-114.6539	89.23864
OA	-7.165058	2.381378	-3.01	0.003	-11.83848	-2.491632
TF	454.3093	197.1237	2.30	0.021	67.45636	841.1623
CI	200.4557	63.51913	3.16	0.002	75.80014	325.1112
imr	-1337.44	346.0289	-3.87	0.000	-2016.517	-658.3624
_cons	-356.1784	346.032	-1.03	0.304	-1035.262	322.9053

Fuente: Elaboración propia

Anexo 48: Test de White para la regresión sobre el dominio Sierra Norte

White's test for Ho: homoskedasticity
against Ha: unrestricted heteroskedasticity

chi2(110) = 286.04
Prob > chi2 = 0.0000

Cameron & Trivedi's decomposition of IM-test

Source	chi2	df	p
Heteroskedasticity	286.04	110	0.0000
Skewness	62.09	14	0.0000
Kurtosis	5.17	1	0.0229
Total	353.30	125	0.0000

Fuente: Elaboración propia

Anexo 49: Prueba de Multicolinealidad para la regresión sobre el dominio Sierra Norte

Variable	VIF	1/VIF
E2	58.75	0.017022
E	52.83	0.018928
ED	21.70	0.046079
ED2	20.17	0.049590
EX	12.02	0.083202
EX2	11.78	0.084892
imr	4.09	0.244665
CI	1.44	0.694345
BD	1.39	0.717506
OA	1.36	0.735534
TF	1.17	0.853479
THT	1.07	0.932318
S	1.05	0.956340
IDS	1.02	0.979435
Mean VIF	13.56	

Fuente: Elaboración propia

Anexo 50: Regresión sobre el dominio Sierra Centro mediante el modelo de Heckman

Linear regression		Number of obs	=	2,621
		F(14, 2606)	=	143.47
		Prob > F	=	0.0000
		R-squared	=	0.4713
		Root MSE	=	724.43

IM	Robust		t	P> t	[95% Conf. Interval]	
	Coef.	Std. Err.				
E	65.77663	7.304507	9.00	0.000	51.45341	80.09986
E2	-.7157287	.090512	-7.91	0.000	-.8932114	-.538246
ED	-74.48948	12.59833	-5.91	0.000	-99.19322	-49.78574
ED2	9.053818	.7085744	12.78	0.000	7.664392	10.44324
EX	28.18577	7.224711	3.90	0.000	14.01901	42.35252
EX2	-.2759843	.2455218	-1.12	0.261	-.7574218	.2054533
IDS	18.35195	31.11584	0.59	0.555	-42.66231	79.3662
THT	8.359487	.8222083	10.17	0.000	6.747239	9.971734
S	275.8432	28.30302	9.75	0.000	220.3446	331.3419
BD	52.30606	29.97246	1.75	0.081	-6.46618	111.0783
OA	1.902171	1.346231	1.41	0.158	-.7376204	4.541962
TF	62.76964	72.76413	0.86	0.388	-79.91171	205.451
CI	261.3199	39.3293	6.64	0.000	184.2001	338.4397
imr	-581.7084	272.7442	-2.13	0.033	-1116.526	-46.89124
_cons	-1359.363	216.5989	-6.28	0.000	-1784.086	-934.6399

Fuente: Elaboración propia

Anexo 51: Test de White para la regresión sobre el dominio Sierra Centro

White's test for H₀: homoskedasticity
against H_a: unrestricted heteroskedasticity

chi2(111) = 425.78
Prob > chi2 = 0.0000

Cameron & Trivedi's decomposition of IM-test

Source	chi2	df	p
Heteroskedasticity	425.78	111	0.0000
Skewness	86.94	14	0.0000
Kurtosis	16.53	1	0.0000
Total	529.24	126	0.0000

Fuente: Elaboración propia

Anexo 52: Prueba de Multicolinealidad para la regresión sobre el dominio Sierra Centro

Variable	VIF	1/VIF
E	53.29	0.018765
E2	49.64	0.020146
ED2	16.40	0.060977
ED	16.28	0.061407
EX	11.22	0.089126
EX2	10.52	0.095051
imr	2.41	0.415557
CI	1.50	0.664996
TF	1.32	0.760144
BD	1.20	0.830821
OA	1.11	0.897880
THT	1.07	0.936505
S	1.06	0.942714
IDS	1.02	0.982631
Mean VIF	12.00	

Fuente: Elaboración propia

Anexo 53: Regresión sobre el dominio Selva mediante el modelo de Heckman

Linear regression	Number of obs	=	3,785
	F(14, 3770)	=	215.09
	Prob > F	=	0.0000
	R-squared	=	0.5231
	Root MSE	=	706.12

IM	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]
E	44.22994	10.96149	4.04	0.000	22.73892 65.72096
E2	-.4017481	.1301144	-3.09	0.002	-.6568495 -.1466467
ED	-94.01343	11.27096	-8.34	0.000	-116.1112 -71.91566
ED2	11.66795	.6397582	18.24	0.000	10.41364 12.92225
EX	20.5469	4.090681	5.02	0.000	12.52673 28.56706
EX2	-.249128	.1282299	-1.94	0.052	-.5005346 .0022786
IDS	7.612413	26.51142	0.29	0.774	-44.3657 59.59053
THT	6.143753	.7346259	8.36	0.000	4.703451 7.584056
S	158.0061	24.16758	6.54	0.000	110.6233 205.3889
BD	28.06604	23.1876	1.21	0.226	-17.39542 73.52751
OA	-.0010144	1.08475	-0.00	0.999	-2.127767 2.125739
TF	225.1044	59.49185	3.78	0.000	108.4651 341.7437
CI	182.6151	30.76609	5.94	0.000	122.2953 242.9349
imr	-812.8751	618.3963	-1.31	0.189	-2025.299 399.5487
_cons	-879.5312	385.2261	-2.28	0.022	-1634.803 -124.2595

Fuente: Elaboración propia

Anexo 54: Test de White para la regresión sobre el dominio Selva

White's test for Ho: homoskedasticity
against Ha: unrestricted heteroskedasticity

chi2(111) = 400.07
Prob > chi2 = 0.0000

Cameron & Trivedi's decomposition of IM-test

Source	chi2	df	p
Heteroskedasticity	400.07	111	0.0000
Skewness	75.29	14	0.0000
Kurtosis	6.28	1	0.0122
Total	481.65	126	0.0000

Fuente: Elaboración propia

Anexo 55: Prueba de Multicolinealidad para la regresión sobre el dominio Selva

Variable	VIF	1/VIF
E	181.34	0.005515
E2	163.09	0.006131
ED	20.21	0.049492
ED2	19.04	0.052520
EX	9.91	0.100861
EX2	9.74	0.102658
imr	6.29	0.158870
CI	1.55	0.644304
BD	1.34	0.746639
TF	1.20	0.830896
S	1.06	0.940636
THT	1.05	0.949963
OA	1.02	0.979378
IDS	1.01	0.994590
Mean VIF	29.85	

Fuente: Elaboración propia

Anexo 56: Regresión sobre el dominio Lima Metropolitana mediante el modelo de Heckman

Linear regression		Number of obs	=	3,733
		F(14, 3718)	=	77.00
		Prob > F	=	0.0000
		R-squared	=	0.3287
		Root MSE	=	1310.6

IM	Robust		t	P> t	[95% Conf. Interval]	
	Coef.	Std. Err.				
E	126.6565	14.83083	8.54	0.000	97.57917	155.7339
E2	-1.437012	.1819839	-7.90	0.000	-1.79381	-1.080213
ED	-249.7775	18.90729	-13.21	0.000	-286.8472	-212.7079
ED2	21.54239	1.185406	18.17	0.000	19.21828	23.8665
EX	45.19986	8.078954	5.59	0.000	29.36025	61.03948
EX2	-.6356424	.2380253	-2.67	0.008	-1.102315	-.1689694
IDS	172.1401	44.98271	3.83	0.000	83.9469	260.3333
THT	7.452979	1.307811	5.70	0.000	4.888882	10.01708
S	265.3428	44.13369	6.01	0.000	178.8142	351.8714
BD	-130.2137	60.48376	-2.15	0.031	-248.7983	-11.62913
OA	16.65703	2.396106	6.95	0.000	11.95922	21.35484
TF	249.4438	47.87284	5.21	0.000	155.5842	343.3034
CI	236.9635	38.87837	6.09	0.000	160.7385	313.1885
imr	3636.742	824.5822	4.41	0.000	2020.064	5253.419
_cons	-3276.756	416.7146	-7.86	0.000	-4093.768	-2459.745

Fuente: Elaboración propia

Anexo 57: Test de White para la regresión sobre el dominio Lima Metropolitana

White's test for H_0 : homoskedasticity
against H_a : unrestricted heteroskedasticity

chi2(111) = 300.06
Prob > chi2 = 0.0000

Cameron & Trivedi's decomposition of IM-test

Source	chi2	df	p
Heteroskedasticity	300.06	111	0.0000
Skewness	84.43	14	0.0000
Kurtosis	10.47	1	0.0012
Total	394.96	126	0.0000

Fuente: Elaboración propia

Anexo 58: Prueba de Multicolinealidad para la regresión sobre el dominio Lima Metropolitana

Variable	VIF	1/VIF
E	140.72	0.007106
E2	140.29	0.007128
ED2	14.60	0.068509
ED	13.54	0.073867
EX	8.44	0.118532
EX2	7.96	0.125688
imr	4.07	0.245654
OA	1.62	0.618524
TF	1.44	0.696766
CI	1.43	0.698585
BD	1.27	0.790274
THT	1.05	0.951365
S	1.03	0.966857
IDS	1.00	0.996726
Mean VIF	24.17	

Fuente: Elaboración propia

Anexo 59: Regresión sobre el dominio Costa Sur mediante el modelo de Heckman

Linear regression		Number of obs	=	1,661
		F(14, 1646)	=	59.92
		Prob > F	=	0.0000
		R-squared	=	0.4011
		Root MSE	=	928.87

IM	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]
E	83.21212	21.74414	3.83	0.000	40.56304 125.8612
E2	-.9213001	.2516486	-3.66	0.000	-1.414885 -.4277149
ED	-88.88605	21.44655	-4.14	0.000	-130.9515 -46.82064
ED2	10.2197	1.178124	8.67	0.000	7.908917 12.53048
EX	33.74651	10.19214	3.31	0.001	13.75558 53.73744
EX2	.0817109	.3250819	0.25	0.802	-.5559067 .7193286
IDS	-67.803	57.83597	-1.17	0.241	-181.2428 45.63684
THT	11.12513	1.416406	7.85	0.000	8.346982 13.90328
S	468.7776	46.94338	9.99	0.000	376.7026 560.8527
BD	-34.65325	57.92361	-0.60	0.550	-148.265 78.95848
OA	.3507923	2.881809	0.12	0.903	-5.301607 6.003191
TF	270.0506	75.45591	3.58	0.000	122.0509 418.0503
CI	218.5097	45.85802	4.76	0.000	128.5635 308.4559
imr	129.4348	637.4058	0.20	0.839	-1120.777 1379.646
_cons	-2030.364	643.2168	-3.16	0.002	-3291.973 -768.7545

Fuente: Elaboración propia

Anexo 60: Test de White para la regresión sobre el dominio Costa Sur

White's test for Ho: homoskedasticity
against Ha: unrestricted heteroskedasticity

chi2(110) = 311.72
Prob > chi2 = 0.0000

Cameron & Trivedi's decomposition of IM-test

Source	chi2	df	p
Heteroskedasticity	311.72	110	0.0000
Skewness	102.74	14	0.0000
Kurtosis	15.90	1	0.0001
Total	430.36	125	0.0000

Fuente: Elaboración propia

Anexo 61: Prueba de Multicolinealidad para la regresión sobre el dominio Costa Sur

Variable	VIF	1/VIF
E	220.21	0.004541
E2	202.34	0.004942
ED	12.82	0.078015
ED2	12.28	0.081416
EX	11.51	0.086887
EX2	11.29	0.088590
imr	7.11	0.140563
OA	1.46	0.686880
CI	1.32	0.758148
TF	1.24	0.805661
BD	1.19	0.839065
THT	1.06	0.944279
S	1.05	0.956348
IDS	1.02	0.984754
Mean VIF	34.71	

Fuente: Elaboración propia

Anexo 62: Regresión sobre el dominio Costa Norte mediante el modelo de Heckman

Linear regression

Number of obs	=	3,972
F(14, 3957)	=	126.67
Prob > F	=	0.0000
R-squared	=	0.4089
Root MSE	=	732.18

IM	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
E	30.58187	9.717775	3.15	0.002	11.52955	49.63418
E2	-.2779988	.120962	-2.30	0.022	-.5151525	-.0408452
ED	-109.6946	10.61756	-10.33	0.000	-130.511	-88.87816
ED2	11.27476	.643631	17.52	0.000	10.01288	12.53664
EX	24.10783	4.320095	5.58	0.000	15.638	32.57765
EX2	-.2655049	.1327334	-2.00	0.046	-.5257371	-.0052726
IDS	7.734643	30.01238	0.26	0.797	-51.10654	66.57582
THT	7.164266	.6582722	10.88	0.000	5.873682	8.454851
S	135.0139	24.66338	5.47	0.000	86.65982	183.3681
BD	69.69234	22.50883	3.10	0.002	25.56234	113.8223
OA	-1.824624	1.499064	-1.22	0.224	-4.763635	1.114387
TF	53.32402	42.97878	1.24	0.215	-30.93862	137.5867
CI	208.0688	28.11285	7.40	0.000	152.9518	263.1858
imr	-1329.602	586.7342	-2.27	0.023	-2479.932	-179.2726
_cons	-466.0145	329.8515	-1.41	0.158	-1112.709	180.6805

Fuente: Elaboración propia

Anexo 63: Test de White para la regresión sobre el dominio Costa Norte

White's test for H_0 : homoskedasticity
against H_a : unrestricted heteroskedasticity

chi2(111) = 357.98
Prob > chi2 = 0.0000

Cameron & Trivedi's decomposition of IM-test

Source	chi2	df	p
Heteroskedasticity	357.98	111	0.0000
Skewness	73.27	14	0.0000
Kurtosis	7.93	1	0.0049
Total	439.17	126	0.0000

Fuente: Elaboración propia

Anexo 64: Prueba de Multicolinealidad para la regresión sobre el dominio Costa Norte

Variable	VIF	1/VIF
E2	156.29	0.006398
E	156.09	0.006407
ED	16.64	0.060100
ED2	15.36	0.065112
EX	9.28	0.107764
EX2	8.82	0.113341
imr	5.01	0.199606
TF	1.49	0.670925
CI	1.48	0.677120
OA	1.43	0.697182
BD	1.22	0.821744
S	1.07	0.931924
THT	1.05	0.950828
IDS	1.01	0.992628
Mean VIF	26.87	

Fuente: Elaboración propia

Anexo 65: Regresión sobre el dominio Costa Centro mediante el modelo de Heckman

Linear regression	Number of obs	=	2,949
	F(14, 2934)	=	95.35
	Prob > F	=	0.0000
	R-squared	=	0.4235
	Root MSE	=	678.11

IM	Robust					[95% Conf. Interval]	
	Coef.	Std. Err.	t	P> t			
E	29.08444	6.289233	4.62	0.000	16.75269	41.4162	
E2	-.2816289	.0798839	-3.53	0.000	-.4382632	-.1249947	
ED	-125.1081	12.13061	-10.31	0.000	-148.8935	-101.3227	
ED2	11.18746	.6932	16.14	0.000	9.82825	12.54667	
EX	30.01081	4.508219	6.66	0.000	21.17122	38.8504	
EX2	-.3881395	.1421041	-2.73	0.006	-.6667733	-.1095057	
IDS	14.79494	26.94192	0.55	0.583	-38.03205	67.62192	
THT	7.137762	.7702979	9.27	0.000	5.627382	8.648141	
S	130.6063	25.91301	5.04	0.000	79.7968	181.4159	
BD	-11.56863	28.63121	-0.40	0.686	-67.70792	44.57066	
OA	3.976357	1.451357	2.74	0.006	1.130576	6.822139	
TF	252.6007	45.97627	5.49	0.000	162.4517	342.7497	
CI	158.904	26.66992	5.96	0.000	106.6103	211.1976	
imr	-1613.307	420.9451	-3.83	0.000	-2438.685	-787.9294	
_cons	-310.9183	202.723	-1.53	0.125	-708.4121	86.57542	

Fuente: Elaboración propia

Anexo 66: Test de White para la regresión sobre el dominio Costa Centro

White's test for Ho: homoskedasticity
against Ha: unrestricted heteroskedasticity

chi2(111) = 425.84
Prob > chi2 = 0.0000

Cameron & Trivedi's decomposition of IM-test

Source	chi2	df	p
Heteroskedasticity	425.84	111	0.0000
Skewness	94.43	14	0.0000
Kurtosis	11.13	1	0.0008
Total	531.41	126	0.0000

Fuente: Elaboración propia

Anexo 67: Prueba de Multicolinealidad para la regresión sobre el dominio Costa Centro

Variable	VIF	1/VIF
E2	57.75	0.017315
E	54.94	0.018203
ED	17.58	0.056869
ED2	15.40	0.064948
EX	9.01	0.110947
EX2	8.58	0.116568
imr	2.63	0.379788
TF	1.34	0.746578
CI	1.32	0.760235
BD	1.15	0.873008
S	1.08	0.926200
OA	1.08	0.929244
THT	1.07	0.936744
IDS	1.01	0.992694
Mean VIF	12.42	

Fuente: Elaboración propia

PARA SUSTENTACION

INFORME DE ORIGINALIDAD

14%

INDICE DE SIMILITUD

14%

FUENTES DE INTERNET

3%

PUBLICACIONES

2%

TRABAJOS DEL
ESTUDIANTE

FUENTES PRIMARIAS

1	www.iadb.org Fuente de Internet	3%
2	hdl.handle.net Fuente de Internet	1%
3	decon.edu.uy Fuente de Internet	1%
4	www.scribd.com Fuente de Internet	<1%
5	www.researchgate.net Fuente de Internet	<1%
6	es.slideshare.net Fuente de Internet	<1%
7	dialnet.unirioja.es Fuente de Internet	<1%
8	docplayer.es Fuente de Internet	<1%
9	repositorio.ulima.edu.pe Fuente de Internet	<1%
