



UNIVERSIDAD NACIONAL DEL ALTIPLANO
FACULTAD DE INGENIERÍA ECONOMICA
ESCUELA PROFESIONAL DE INGENIERÍA ECONÓMICA



**INFLUENCIA DE LA EDUCACIÓN Y LA CREACIÓN DE NUEVAS
EMPRESAS SOBRE EL EMPLEO INFORMAL Y DIFERENCIA DE
INGRESOS ENTRE OCUPACIONES INFORMALES Y
FORMALES EN PERÚ, PERIODO 2022**

TESIS

PRESENTADA POR:

Bach. PEDRO FERNANDO FLORES TACURI

PARA OPTAR EL TÍTULO PROFESIONAL DE:

INGENIERO ECONOMISTA

PUNO - PERÚ

2024



Reporte de similitud

NOMBRE DEL TRABAJO

**Influencia de la educación en el empleo i
nformal y diferencia de ingresos entre oc
upaciones informal**

AUTOR

Pedro Fernando Flores Tacuri

RECuento DE PALABRAS

20082 Words

RECuento DE CARACTERES

107771 Characters

RECuento DE PÁGINAS

101 Pages

TAMAÑO DEL ARCHIVO

1.7MB

FECHA DE ENTREGA

May 15, 2024 5:15 PM GMT-5

FECHA DEL INFORME

May 15, 2024 5:19 PM GMT-5

● 12% de similitud general

El total combinado de todas las coincidencias, incluidas las fuentes superpuestas, para cada base de datos.

- 10% Base de datos de Internet
- Base de datos de Crossref
- 8% Base de datos de trabajos entregados
- 3% Base de datos de publicaciones
- Base de datos de contenido publicado de Crossref

● Excluir del Reporte de Similitud

- Material bibliográfico
- Material citado
- Bloques de texto excluidos manualmente
- Material citado
- Coincidencia baja (menos de 10 palabras)




Dr. Sabino Edgar Mamani Choque
Director de la Unidad de Investigación - FIE
UNA - PUNO


Pedro Fernando Flores Tacuri

Resumen



DEDICATORIA

Quiero expresar mi gratitud a mis abuelos, Leonardo y Filomena, y a mis padres, Martin, Julia y Vidal, por haberme educado con principios, esfuerzo y dignidad. Gracias a ustedes he alcanzado muchos de mis objetivos, entre ellos este.

A mis hermanas Dora, Dayana, Iris, Yaneth y Valeri les agradezco por ser una parte fundamental de mi vida y por motivarme a seguir creciendo y mejorando.

Pedro Fernando Flores Tacuri



AGRADECIMIENTOS

Expreso mi gratitud a Dios por guiarme en mi trayectoria vital, por concederme salud y vigor para avanzar y superar los obstáculos. A mi padre, por transmitirme valores de sencillez, humildad y perseverancia, por estar a mi lado y respaldarme siempre. A mi madre, por confiar en mí y acompañarme en cada instante, por sus sabios consejos y su amor incondicional. Gracias a los dos por ser mi sostén y mi guía.

Agradezco a la Universidad Nacional del Altiplano, a la Facultad de Ingeniería Económica, y especialmente a los docentes por todo lo que me instruyeron y me facilitaron el aprendizaje durante mi formación profesional.

Del mismo modo, reconozco a mis jurados por sus sugerencias, apoyo y dedicación para la culminación de mi tesis.

A mi asesor el Dr. Roberto Arpi Mayta, por todo su apoyo, dirección y atención para la ejecución y finalización de mi tesis.

Pedro Fernando Flores Tacuri



ÍNDICE GENERAL

	Pág.
DEDICATORIA	
AGRADECIMIENTOS	
ÍNDICE GENERAL	
ÍNDICE DE TABLAS	
ÍNDICE DE FIGURAS	
ÍNDICE DE ANEXOS	
ACRÓNIMOS	
RESUMEN	13
ABSTRACT.....	14
CAPÍTULO I	
INTRODUCCIÓN	
1.1. PLANTEAMIENTO DEL PROBLEMA.....	17
1.2. JUSTIFICACIÓN DEL ESTUDIO.....	19
1.3. OBJETIVOS DE INVESTIGACIÓN.....	20
1.3.1. Objetivo general	20
1.3.2. Objetivo específicos	20
CAPÍTULO II	
REVISIÓN DE LITERATURA	
2.1. ANTECEDENTES	21
2.1.1. A nivel Internacional.....	21
2.1.2. A nivel Nacional	25
2.2. MARCO TEÓRICO	32
2.2.1. Teoría Neoclásica del Mercado de trabajo.....	32



2.2.2. El Modelo Teórico Ocio Consumo	35
2.2.3. Teoría de segmentación, en relación a la brecha de ingreso	37
2.3. MARCO CONCEPTUAL	40
2.4. MÉTODOS EMPÍRICOS PARA LA ESTIMACIÓN	42
2.4.1. Metodología econométrica Logit y Probit	42
2.4.2. Metodología para la estimación de la Brecha de Ingresos	45
2.5. HIPÓTESIS DE LA INVESTIGACIÓN	50
CAPÍTULO III	
MATERIALES Y MÉTODOS	
3.1. UBICACIÓN GEOGRÁFICA DEL ESTUDIO.....	52
3.2. MÉTODO Y TIPO DE INVESTIGACIÓN	53
3.3. PROCEDENCIA DEL MATERIAL UTILIZADO	53
3.4. POBLACIÓN Y MUESTRA DE ESTUDIO	53
3.5. PROCEDIMIENTO	54
3.6. VARIABLES	60
CAPÍTULO IV	
RESULTADOS Y DISCUSIÓN	
4.1. RESULTADOS	61
4.1.1. Análisis descriptivo de las variables	61
4.1.2. Resultados del análisis econométrico.....	74
4.2. DISCUSIÓN.....	82
V. CONCLUSIONES.....	86
VI RECOMENDACIONES	88
VII. REFERENCIAS BIBLIOGRÁFICAS.....	89
ANEXOS.....	95



ÁREA: Ciencias Económicas Empresariales

TEMA: Políticas Públicas

FECHA DE SUSTENTACIÓN: 21 de mayo de 2024



ÍNDICE DE TABLAS

Tabla 1	Organización de variables.....	60
Tabla 2	Brecha del Ingreso Promedio y Número de Casos por Regiones Naturales	64
Tabla 3	Ingreso Promedio y Número de Casos por Ámbito.....	66
Tabla 4	Ingreso Promedio y Número de Casos por Sexo	67
Tabla 5	Ingreso Promedio por Nivel Educativo.....	70
Tabla 6	Número de Casos por Nivel Educativo.....	72
Tabla 7	Criterios estadísticos y valores para seleccionar el modelo más adecuado .	74
Tabla 8	Resultados del <i>Modelo Logit y Probit</i> estimado.....	75
Tabla 9	Efectos marginales del modelo estimado.....	77
Tabla 10	Resultados obtenidos de la <i>metodología Oaxaca - Blinder</i>	79
Tabla 11	Resultados obtenidos de la <i>Metodología Ñopo</i>	80



ÍNDICE DE FIGURAS

Figura 1	Nivel de equilibrio	37
Figura 2	Mapa del Perú.....	52
Figura 3	Porcentaje de formalidad e informalidad	62
Figura 4	Mapa del Perú características del empleo a nivel de ingreso, cantidad y empresa.....	63



ÍNDICE DE ANEXOS

ANEXOS A Fuente de Obtención de datos.....	95
ANEXOS B Resultado de estimación logit.....	96
ANEXOS C Resultados de efectos marginales del modelo logit.....	96
ANEXOS D Resultados estimación modelo probit	97
ANEXOS E Resultado de estimación efectos marginales modelo probit.....	97
ANEXOS F Resultados de estimación por método de Oaxaca - Blinder	98
ANEXOS G Resultados de estimación método de Ñopo	99



ACRÓNIMOS

OIT:	Organización Internacional de Trabajo
PEA:	Población Económicamente activa
PET:	Población en edad de trabajar
INEI:	Instituto Nacional de Estadística e Informática
ENAHO:	Encuesta Nacional de Hogares.
SIRTOD:	Sistema de Información Regional para la Toma de Decisiones
ENOE:	Encuesta Nacional de Ocupación y Empleo - México
INEGI:	Instituto Nacional de Estadística, Geografía e Informática - México
ECETSS:	Encuesta Nacional a Trabajadores sobre Condiciones de Empleo, Trabajo, Salud y Seguridad - Argentina.
SRT:	Superintendencia de Riesgos del trabajo – Argentina.
FDA:	Función de distribución acumulativa.
INGM:	Ingreso laboral mensual (en logaritmos) deflactado.
INGH:	Ingreso laboral por hora (en logaritmos) deflactado.
OCUPINF:	Situación de informalidad (Ocup. Principal) (Informal o Formal).
EDU:	Años de educación alcanzados.
EXP:	Experiencia potencial.



EXP2:	Experiencia potencial al cuadrado.
NEMP:	Número de nuevas empresas.
TEMP:	Número de trabajadores en la empresa.



RESUMEN

La presente investigación se enfoca en analizar la variable educación y la creación de nuevas empresas, con el propósito de determinar su impacto sobre el empleo informal y las disparidades de ingresos entre ocupaciones informales y formales en Perú durante el año 2022. La problemática de la informalidad laboral es una cuestión generalizada a nivel global, como señala la OIT (2020), indicando que el 61.2% de los trabajadores a nivel mundial están empleados de manera informal. En el caso peruano, este fenómeno tiene repercusiones negativas en la economía, la sociedad y la calidad de vida, alcanzando un 72% de la Población Económicamente Activa ocupada en 2019, según datos del Instituto Nacional de Estadística e Informática (2020). Para el primer objetivo, se aplicaron los modelos logit y probit, y para el segundo objetivo, se emplearon los métodos de descomposición Oaxaca–Blinder y la Ñopo. Estas herramientas permitieron obtener resultados significativos. La educación disminuye la probabilidad de informalidad en un 4.04%, la experiencia potencial reduce esta probabilidad en un 1.06%, y la variable de género muestra un aumento del 2.55% en la probabilidad para los hombres. Además, el tamaño de la variable número de trabajadores en la empresa reduce en un 10.50%, según el modelo probit. En relación con la brecha salarial, según el método Oaxaca–Blinder, la diferencia es de 0.97, unidades logarítmicas entre formales e informales. De acuerdo con el método de Ñopo, se observa que la diferencia logarítmica alcanza a 0.872 entre trabajadores formales e informales.

Palabras clave: empleo Informal, educación, descomposición Oaxaca – Blinder, nuevas empresas, tamaño de empresa.



ABSTRACT

The present research focuses on analyzing the relationship between the education variable and the creation of new businesses, with the purpose of determining its impact on informal employment and income disparities between informal and formal occupations in Peru during the year 2022. The problem of labor informality is a widespread issue at the global level, as pointed out by the OIT (2020), indicating that 61.2% of workers worldwide are employed informally. In the Peruvian case, this phenomenon has negative repercussions on the economy, society and quality of life, reaching 72% of the Economically Active Population employed in 2019, according to data from the National Institute of Statistics and Informatics (2020). The methodology used to address the first objective, through logit and probit, as well as the Oaxaca–Blinder decomposition technique and the Ñopo for the second objective, revealed significant results. Education decreases the probability of informality by 4.04%, potential experience reduces this probability by 1.06%, and the gender variable shows an increase of 2.55% in the probability for men. In addition, the number of businesses in Peru decreases by 0.0006%, and the size of the variable number of workers in the company reduces by 10.50%, according to the probit model. In relation to the wage gap, according to the Oaxaca–Blinder method, formal workers earn 0.97 logarithm units more than informal workers. According to the Ñopo method, it is observed that the labor income of formal workers is 0.872 logarithm units more than informal workers.

Keywords: informal employment, education, Oaxaca – Blinder decomposition, new businesses, firm size.



CAPÍTULO I

INTRODUCCIÓN

En la actualidad, el fenómeno del empleo informal representa un desafío significativo a nivel global, afectando a más del 60% de los trabajadores en todo el mundo OIT (2020). En este contexto, Perú no es una excepción, y a pesar de los avances económicos reportados por el Ministerio de Economía y Finanzas (MEF 2023) y el incremento en el número de empresas según el Instituto Nacional de Estadística e Informática (INEI 2022), la tasa de empleo informal se mantiene alarmantemente alta.

La presente investigación se enfoca en comprender la complejidad de los factores que contribuyen a la persistencia de la informalidad laboral en Perú durante el periodo 2022. En particular, se busca explorar la influencia de la educación y la creación de nuevas empresas como variables clave en este fenómeno, en consonancia con el objetivo 8 de la Agenda 2030 de la Organización Internacional del Trabajo (OIT 2017), centrado en el trabajo decente y el crecimiento económico.

La literatura revisada destaca la importancia de abordar la informalidad desde una perspectiva integral, considerando no solo aspectos económicos sino también socioculturales. Estudios previos como los de de Ovando et al. (2021), Paz (2021), Casarreal y Cruz (2020), y otros, coinciden en la relevancia de la educación y la creación de empresas como factores clave para comprender la informalidad laboral. En este sentido, la presente investigación se propone analizar la relación entre estas variables y su impacto en la ocupación informal y en la disparidad de renta entre trabajos informales y formales.



La situación en Perú, a pesar de su crecimiento económico, refleja una realidad preocupante, con una tasa de empleo informal del 76.1% entre abril de 2021 y marzo de 2022, según datos del INEI (2022). Esta situación no solo vulnera los derechos laborales, sino que también contribuye a la desigualdad económica, ya que, según estudios como el de Machaca (2019) y Maurizio (2014), los trabajadores informales suelen percibir ingresos inferiores a sus pares formales.

En este contexto, las incógnitas que guiarán la investigación se centran en la comprensión de cómo la educación y la creación de nuevas empresas influyen en la ocupación informal y en qué medida explican la disparidad de ingresos entre trabajo informales y formales en el contexto peruano durante el año 2022.

La presente investigación se plantea como una contribución significativa para entender y abordar la problemática del empleo informal en Perú, proporcionando conocimientos que puedan orientar políticas educativas y estrategias de formalización laboral. Además, busca responder a la necesidad de explorar estas dinámicas en el contexto específico peruano, enriqueciendo la literatura existente sobre el empleo informal y su relación con la educación y la creación de empresas. Para tal objetivo, comprende los siguientes capítulos

En el primer capítulo contiene la introducción, donde se plantea el problema de investigación, se justifica su relevancia y se definen los objetivos que se pretenden alcanzar.

En el segundo capítulo se revisan los antecedentes de investigación y el marco teórico que sustenta el estudio.



En el tercer capítulo se explica los materiales y métodos que se utilizaron en este estudio, especificando la localización geográfica, el método, el tipo de investigación, el origen del material usado, la población y la muestra, el procedimiento y las variables.

El capítulo IV presenta los resultados y discusión; se realiza un análisis descriptivo de las variables, luego se presentan los resultados de la regresión econométrica, este capítulo sigue el enfoque de los objetivos; para finalizar con las conclusiones y las recomendaciones.

1.1. PLANTEAMIENTO DEL PROBLEMA

Según la OIT (2020), el empleo informal abarca un 61,2 por ciento de los trabajadores del mundo, de los cuales el 52 por ciento pertenecen al sector informal, el 6,7 por ciento al sector formal y el 2,5 por ciento a los hogares. Por esta razón, es urgente entender este fenómeno, para que a partir de ello podamos enfrentar los desafíos que implica generar empleo y condiciones laborales dignas, enfatizando el rol de las políticas educativas y las estrategias de formalización laboral para reducir el empleo informal, tal como lo establece el objetivo 8 de la agenda al 2030 de la Organización Internacional del Trabajo (2017), llamado trabajo decente y crecimiento económico. Asimismo teniendo en cuenta los estudios realizados en el ámbito internacional como los de Ovando et al. (2021), Paz (2021), Casarreal y Cruz (2020), Carvajal et al. (2017), Ramírez et al. (2015), Hernández et al. (2014) y Maurizio (2014) quienes investigan el empleo informal en diversos contextos, coincidiendo en que la educación y la creación de empresas son factores clave para entender la informalidad laboral.

La situación laboral en Perú está influida por aspectos económicos y socioculturales, que determinan de forma compleja la creación de nuevas empresas, la educación y la informalidad laboral. A pesar de que el MEF (2023) indica que el país ha



tenido un importante crecimiento económico en la última década, salvo en el 2020, y que el INEI (2022) reporta un aumento del stock de empresas, la informalidad laboral se ha mantenido a lo largo de los años. Según el INEI (2022), entre abril del 2021 y marzo del 2022, la tasa de empleo informal fue de 76,1%, mientras que la tasa de empleo formal fue de 23,9%. Esta realidad es alarmante ya que no solo implica una violación de los derechos laborales de los trabajadores, sino también una brecha de ingresos entre los empleos formales e informales. De acuerdo con la literatura revisada, como lo señalado por Machaca (2019) y Maurizio (2014), los trabajadores informales suelen percibir menos ingresos que los trabajadores formales, lo que genera desigualdad económica. Además, considerando las teorías mencionadas por Becker (1975), Krugman y Wells (2009), Nicholson (2008) y Blanchard et al. (2012), se puede afirmar que el mercado laboral es un fenómeno complejo y que uno de los factores asociados a la informalidad es el nivel educativo.

En ese sentido, es importante analizar la influencia de la educación y la creación de nuevas empresas formales en la informalidad laboral y en la brecha de ingresos entre empleos formales e informales. Es conocido que una mayor educación puede mejorar las oportunidades de empleo y los ingresos de los trabajadores, pero ¿en qué medida esto se cumple en el contexto peruano? Bajo este escenario se presenta las siguientes incógnitas:

1.1.1. problema general

¿Cuál es la influencia de la variable Educación y la creación de nuevas empresas sobre el empleo informal y en la diferencia de ingresos entre ocupaciones informales y formales en Perú, periodo 2022?

1.1.2. problemas específicos



- ¿Cómo influye la variable educación y la creación de nuevas empresas sobre la ocupación informal en Perú, periodo 2022?
- ¿En qué medida explica la variable educación y la creación de nuevas empresas a la diferencia de ingresos entre ocupaciones informales y formales en Perú, periodo 2022?

1.2. JUSTIFICACIÓN DEL ESTUDIO

El 2022 representa un periodo posterior a la pandemia de COVID – 19, lo que permite analizar la situación laboral y económica del Perú en un contexto de recuperación económica. El problema de la informalidad es un problema común en muchos países así como lo mencionado por la OIT (2020) que precisa que un 61.2 por ciento de los trabajadores del mundo pertenecen al empleo informal. En el Perú, este tiene impactos negativos en la economía, la sociedad y la vida de las personas, Según los datos de Instituto Nacional de Estadística e Informática (2020), la informalidad laboral en el 2019 alcanza un 72% de la PEA ocupada. Además, según la literatura es que existe una brecha significativa en el ingreso entre estos empleos. La educación se ha relacionado con la posibilidad de conseguir empleos formales y mejores ingresos, pero se desconoce el impacto que tiene en el 2022 en el mercado laboral peruano. Por lo tanto, entender cómo influye la educación y la creación de nuevas empresas en el empleo informal y la brecha de ingresos, puede tener implicaciones importantes para las políticas públicas, además de proporcionar información empírica, lo que sería útil para los responsables de formular políticas públicas, académicos y otros interesados en el desarrollo económico y social del país.



1.3. OBJETIVOS DE INVESTIGACIÓN

1.3.1. Objetivo general

Analizar la variable educación y la creación de nuevas empresas y determinar su influencia en el empleo informal y en la diferencia de ingresos entre ocupaciones informales y formales en Perú, periodo 2022.

1.3.2. Objetivo específicos

- Analizar la influencia de la variable educación y la creación de nuevas empresas sobre la ocupación informal en Perú, periodo 2022.
- Determinar en qué medida explica la variable educación y la creación de nuevas empresas en la diferencia de ingresos entre ocupaciones informales y formales en Perú, periodo 2022.



CAPÍTULO II

REVISIÓN DE LITERATURA

2.1. ANTECEDENTES

2.1.1. A nivel Internacional

Billorou (2021) analiza la diferencia de ingresos entre mujeres rurales y urbanas en Colombia en 2019, usando una muestra de 24,777,058 personas. Aplica el modelo de Oaxaca-Blinder para descomponer la brecha. Las variables que considera son la educación, experiencia, experiencia al cuadrado, horas trabajadas, departamento, actividad, sector y ocupación. Encuentra que las horas trabajadas son clave, con un 52% de probabilidad de aumento de ingreso de la población urbana y un 49% en la población rural. También, que un año más de estudio aumenta el 9% en el ingreso de la población urbana y el 7% en la población rural. Aunque el modelo explica el 59.1% de la brecha con estas variables, queda un 29% sin explicar y esta se relaciona con la locación de la mujer (Billorou 2021).

Ovando et al. (2021) analizan la situación del empleo informal en México en dos años clave, 2005 y 2020. Se consideran variables como sexo, edad, estado civil, escolaridad, localidad, tamaño de empresa, sector de la actividad económica y horas de trabajo, con base en la ENOE del INEGI. Se utiliza un modelo de regresión logística binaria para contrastar las razones de probabilidad entre grupos y explicar las diferencias en los efectos entre los periodos. Los resultados indican que, en 2005, la probabilidad de que un hombre tenga empleo informal es 1.047:1 veces mayor que la de una mujer, pero en 2020, esta relación se reduce a 0.966:1,



lo que significa una mayor presencia de las mujeres en el empleo informal en el último periodo (Ovando et al. 2021).

Paz (2021) tiene como objetivo examinar y comprender los factores que determinan la disparidad de ingresos entre trabajadores formales e informales en Argentina. La investigación se centra en cuatro categorías ocupacionales: asalariado, doméstico, empleador y autónomo. Los datos provienen de la Encuesta Nacional a Trabajadores sobre Condiciones de Empleo, Trabajo, Salud y Seguridad (ECETSS), realizada por la Superintendencia de Riesgos del Trabajo (SRT). La metodología empleada se apoya en el enfoque de Chernozhukov para analizar las desigualdades salariales a lo largo de la distribución. Los resultados indican que los ingresos de los ocupados en el sector formal son, en promedio, entre 2.5 (146 %) y 1.5 (55 %) veces más elevados que los del sector informal. Asimismo, se observa que la diferencia salarial entre los ingresos de los ocupados en el sector formal y el promedio es menor que la que existe entre los ingresos de los ocupados en el sector informal y el promedio (Paz 2021).

Casarreal y Cruz (2020) analizan la demanda del empleo informal en América Latina, basándose en diferentes teorías, como el mercado de trabajo neoclásico, el modelo de crecimiento de Lewis y la rigidez del mercado laboral, y en las políticas que afectan a las empresas. Usando datos de 1995-2016 de países de Latinoamérica tales como Perú, Bolivia, Argentina, Chile, Colombia, México, Venezuela y Brasil, elegidos por su calidad de datos, los autores muestran que la reducción del empleo informal está vinculada con el crecimiento del sector manufacturero en la economía total. Esto significa un aumento en la productividad y, por lo tanto, una mayor demanda de trabajadores. La conclusión final resalta la importancia de implementar políticas de re-industrialización para disminuir el



empleo informal, enfatizando el rol del gobierno mediante el gasto público en inversiones. Esta propuesta es relevante para América Latina, donde la región ha buscado el crecimiento y desarrollo económico mediante la globalización, sin éxito hasta ahora (Casarreal y Cruz 2020).

Carvajal et al. (2017) tienen como objetivo examinar cómo influyen las condiciones socioeconómicas en la precariedad y la informalidad laboral en la zona metropolitana de Bucaramanga (ZMB), Colombia. Para ello, se realiza un análisis descriptivo y se estima un modelo biprobit con datos de la Gran Encuesta Integrada de Hogares de 2015. Los resultados indican que el 16,52% de los trabajadores en el AMB se encuentran en situación de subempleo e informalidad laboral simultáneamente. Entre los factores que reducen la probabilidad de pertenecer a este grupo se destacan el nivel educativo, el género masculino y la edad. Por el contrario, aspectos como el tiempo entre empleos y la actividad en sectores económicos secundarios o terciarios aumentan la probabilidad conjunta de tener un empleo precario e informal. Este estudio contribuye a una comprensión integral de los determinantes socioeconómicos de la precariedad y la informalidad laboral en el AMB (Carvajal et al. 2017).

Ramírez et al. (2015) tienen como objetivo identificar los factores que influyen en la permanencia en el sector informal de la economía en Colombia entre 2008 y 2012. La pregunta que guía la investigación es cómo las políticas públicas dirigidas a la población pobre inciden en las decisiones individuales de participar en el mercado laboral formal o informal. A partir de los datos de la Gran Encuesta Integrada de Hogares (GEIH) del DANE, se estima un modelo Probit, donde la variable dependiente es la informalidad y las variables independientes son características como el género, la edad, el nivel educativo y las horas



trabajadas por semana. Los resultados indican que las mujeres tienen una probabilidad de estar en el sector informal de un 11.73%, y que cada diez años de edad disminuyen las probabilidades de informalidad en un 6.9%. Además, un mayor nivel educativo reduce en un 10% la probabilidad de informalidad (Ramírez et al. 2015).

Hernández et al. (2014) en Colombia, exploran la complejidad de la informalidad laboral, destacando el desempleo, el nivel educativo y el género como factores sociodemográficos que influyen en la elección de trabajar en el sector informal. También consideran aspectos económicos como la globalización, el desarrollo industrial, el ciclo económico y el tipo de cambio real, que condicionan las oportunidades y los incentivos para participar en el mercado laboral formal. Enfatizan la importancia de la regulación del mercado laboral, que puede generar distorsiones y costos que favorecen la informalidad, convirtiéndola en una opción de ingresos ante la falta de empleo formal. Este estudio proporciona una síntesis de los principales hallazgos sobre la informalidad y sus determinantes, subrayando que este problema es el resultado de una combinación de factores económicos, sociodemográficos e institucionales (Hernández et al. 2014).

Maurizio (2014) examina en profundidad la situación y los rasgos de la informalidad laboral en América Latina, así como las diferencias salariales que se derivan de esta realidad. Utiliza micro datos de encuestas a hogares de 11 países de la región, elegidos por la disponibilidad de información y un tamaño de muestra representativo. Para calcular las diferencias salariales, usa tres métodos distintos: las Ecuaciones Mincerianas con el Método de Heckman en dos fases, la Regresión por Cuantiles y el método de Descomposición de Oaxaca-Blinder. La investigación divide los países en dos grupos, resaltando la gran heterogeneidad



en la composición ocupacional. Los resultados muestran diferencias salariales promedio que varían alrededor del 34% en Argentina y Ecuador, el 28% en Uruguay y cerca del 20% en los demás casos, con un 24.6% en el caso peruano. Este estudio afirma que la informalidad y la desigualdad salarial son características inherentes al mercado laboral latinoamericano, contribuyendo a niveles importantes de precariedad y salarios bajos (Maurizio 2014).

Carrillo (2004) estudia la brecha salarial entre el sector formal e informal en Ecuador. Encuentra que aspectos sociodemográficos, como el desempleo, la educación y el género, afectan la decisión de trabajar en el sector informal. También analiza variables económicas, como la globalización, el desarrollo industrial, el ciclo económico y el tipo de cambio real, que impactan en las posibilidades e incentivos para participar en el mercado laboral formal. Además, enfatiza el rol de la regulación del mercado laboral, que puede crear distorsiones y costos que favorecen la informalidad. Los hallazgos muestran que la informalidad es una alternativa de ingresos ante la escasez de empleo formal, y los trabajadores del sector público, en promedio, ganan un 18% más que los del sector privado. Indica, también que las diferencias salariales son más marcadas para los que ganan menos del promedio, disminuyendo entre los que ganan más (Carrillo 2004).

2.1.2. A nivel Nacional

Adanaqué (2022) analiza cómo la educación influye en la brecha salarial, usando solo datos peruanos de fuentes relevantes como los módulos 3 y 5 de la ENAHO y el SIRTOD. Para dar conclusiones sólidas a nivel regional y de interés nacional, usa un modelo de panel de datos con efectos fijos que cubre todos los



departamentos peruanos entre 2008 y 2019. Los hallazgos muestran que el aumento del nivel educativo promedio en los departamentos se relaciona con una reducción de la brecha salarial. El análisis muestra que la desigualdad de los ingresos salariales no se ve afectada de manera significativa por la brecha educativa. Además, el estudio respalda el uso del modelo de panel de datos con efectos fijos frente a los modelos de efectos aleatorios y MCO (Adanaqué 2022).

Solis (2021) examina la informalidad laboral desde el lado de la oferta, enfocándose en los factores clave que afectan la informalidad laboral en la provincia de Lima en 2018. El estudio busca determinar los factores más relevantes, como el nivel educativo, sexo, estado civil y edad, usando datos de la Encuesta Nacional de Hogares (ENAH) del INEI. La metodología empleada usa modelos econométricos Logit y Probit para estimar la probabilidad de ser trabajador informal en Lima en 2018 y para obtener las probabilidades esperadas de estar en la informalidad. Los hallazgos muestran que, al incrementar el nivel educativo, la probabilidad de informalidad se reduce en un 20.09%. Asimismo, el sexo femenino, el aumento de la edad y otros factores se relacionan con aumentos en la probabilidad de informalidad. Se resalta que los hombres entre 23 y 24 años, con educación universitaria completa y estado civil soltero, tienen la menor probabilidad (4.24%) de ser informales. Por el contrario, las mujeres de 65 años o más, con educación secundaria incompleta o menor y estado civil casada o conviviente, tienen la mayor probabilidad (89.01%) de ser trabajadoras informales, destacando la complejidad y diversidad de los factores que influyen en la informalidad laboral en Lima (Solis 2021).

Esparta y Rivera (2020) examinan la dinámica de las diferencias salariales entre el sector formal e informal en el mercado de trabajo peruano, enfocándose



en la discriminación o segmentación que contribuye a esta desigualdad. Usando datos de la ENAHO del INEI para el período 2007-2019, el análisis incorpora diversos factores, como educación, capacitación, acceso a servicios públicos, localización geográfica, costos de movilidad laboral, información incompleta e imperfecta, salario mínimo, entre otros. La metodología utilizada comprende tres enfoques de descomposición: el Método de descomposición paramétrica de Oaxaca y Blinder, el Método de descomposición no paramétrica de Ñopo, y el Método de descomposición por cuantiles de Machado y Mata. Los hallazgos señalan que las características no observables predominan, con el 38.1% de la diferencia salarial atribuido a características observables en 2019 y el 69.2% sin explicar. Ñopo muestra que las características no observables representan el 40.7%, mientras que Machado y Mata resaltan una diferencia salarial de más del 100%, siendo más acentuada en trabajadores de ingresos más bajos, explicada principalmente por características no observables asociadas a la discriminación (Esparta y Rivera 2020).

Ramos (2020) exploró las variables vinculadas al empleo informal en el departamento de Puno durante el período 2008-2017, fundamentándose en diversas teorías del mercado laboral, como el Sistema Clásico, el Sistema Neoclásico, el Sistema Keynesiano y Poskeynesiano, así como el Sistema Moderno. Adoptó también enfoques estructuralistas e institucionalistas para abordar la informalidad. Utilizando el método inductivo y explicativo, se basó en datos de la Encuesta Nacional de Hogares (ENAHO) correspondientes a esos años. La metodología empleada fue la estimación econométrica Probit, con la variable dependiente siendo el empleo informal y variables independientes como educación, jefe de hogar, género, edad y tamaño de empresa. Los resultados



revelaron que la educación mostró una relación negativa a lo largo del período, reduciendo la probabilidad de participación en el mercado laboral en un 3%. De manera similar, la edad y el tamaño de la empresa disminuyeron la probabilidad de participación laboral en un 2.8% y 4%, respectivamente. Las variables de género y ser jefe de hogar resultaron significativas solo en ciertos años (Ramos 2020).

MACHACA (2019) tiene como objetivo examinar los determinantes de la informalidad y la brecha salarial entre los empleados formales e informales en el Perú, utilizando datos de la Encuesta Nacional de Hogares de 2017. El marco teórico se basa en la teoría neoclásica y la teoría de segmentación del mercado de trabajo. La teoría neoclásica sostiene que la informalidad se origina por la baja productividad laboral asociada a niveles educativos y experiencias laborales deficientes. Por su parte, la teoría de segmentación propone que existen barreras que impiden el tránsito entre sectores, incluso cuando los trabajadores tienen características individuales similares, lo que se denomina discriminación. La metodología empleada es hipotético-deductiva, aplicando el modelo econométrico de Mincer con la corrección de Heckman para el sesgo de selección. La brecha salarial se analiza con la metodología de Oaxaca-Blinder. Los resultados indican que la informalidad está influida por factores como el nivel educativo (-6%), la experiencia laboral (-0.5%), el género y el tamaño de la empresa (78%). Asimismo, la brecha salarial, en logaritmos, es de 0.869, de la cual el 30% se explica por ventajas en las características individuales y el 54.1% se debe a la discriminación (MACHACA 2019).

Sánchez y Chafloque (2019) utilizando la ENAHO del INEI, estimaron la informalidad laboral en el Perú en 2014. Su muestra incluyó a 16,157,315



ocupados a nivel nacional, y aplicaron análisis factorial para clasificarlos en tres grupos: informalidad plena, informalidad en empleo formal e informalidad en trabajadores independientes. El comercio ambulatorio fue un caso especial de estudio. Hallaron que 9,975,633 personas eran informales, 4,454,164 tenían empleo formal pleno y 1,727,518 eran informales en empleo formal. Contrariamente a lo que se suele pensar, la informalidad no siempre implica pobreza o precariedad, pues hay actividades informales con mayores ingresos que el empleo formal. Además, la representación política territorial influye en las condiciones laborales formales e informales. La investigación aportó una cuantificación y una georreferenciación de la informalidad laboral a diferentes niveles (macro zonas, zonas ENAHO y provincias), y generó un mapa nacional que puede ser útil para diversos actores interesados en comprender mejor la realidad laboral peruana (Sánchez y Chafloque 2019).

Alvarado (2018) examina la informalidad laboral en Perú desde la perspectiva de la teoría de mercados laborales multisegmentados. La investigación revisa diferentes teorías, como la de segmentación laboral y los enfoques de informalidad laboral, para estimar el porcentaje de trabajadores informales involuntarios y analizar factores pecuniarios y no pecuniarios entre 2012 y 2016. Los resultados, obtenidos mediante un modelo que ajusta supuestos previos, confirman la existencia de mercados laborales multisegmentados, identificando que entre el 9% y el 13% de los asalariados urbanos del sector privado son trabajadores informales involuntarios. Se concluye que la educación superior reduce la probabilidad de informalidad, mientras que la tenencia de seguros diferentes al del trabajo aumenta dicha probabilidad. Además, las responsabilidades familiares disminuyen la probabilidad de ser asalariado



informal. Este estudio aporta a desmitificar la relación entre informalidad laboral y precariedad, destacando la existencia de actividades informales que generan ingresos superiores al empleo formal, y señala que la informalidad no es exclusivamente marginal, sino que varía geográfica y territorialmente en el Perú (Alvarado 2018).

Pedraza (2016) analiza los determinantes de la informalidad laboral en el Departamento del Cusco en el año 2014, a partir de distintas teorías del mercado de trabajo y empleando la Encuesta Nacional de Hogares (ENAHO). El método empleado es un modelo probit que utiliza como variables explicativas el nivel de educación, el estado civil, la edad y el tamaño de la empresa, y como variable de interés la informalidad laboral. Los resultados indican que la educación, el estado civil y la edad tienen una relación inversa con la informalidad laboral, es decir, a mayor nivel de estos factores, menor es la probabilidad de ser informal. Por otro lado, el tamaño de la empresa tiene una relación directa con la informalidad, siendo más alta en las microempresas. Estos resultados ponen de manifiesto la relevancia de abordar la informalidad considerando tanto factores individuales como condiciones del mercado de trabajo en el Departamento del Cusco en ese período (Pedraza 2016).

Rentería (2015) examina las diferencias en los ingresos laborales en las ciudades peruanas, con especial atención a la economía informal. Usando datos de la Encuesta Nacional de Hogares (ENAHO) del Instituto Nacional de Estadística e Informática (INEI), que se recoge desde 2004, el estudio usa una técnica de emparejamiento para descomponer las brechas salariales entre grupos distintos según su nivel de informalidad. Esta metodología, basada en la propuesta de Ñopo (2004), trata de analizar factores explicables y no explicables en las



brechas salariales. Los resultados muestran que los asalariados con empleos formales tienen casi ninguna posibilidad de cambiar a empleo asalariado en el sector informal, pero podrían pasar a un empleo informal dentro de empresas formales. En promedio, el componente no explicado de las brechas de ingresos entre formales e informales de la misma categoría de ocupación llega al 35%. Además, se resalta la importante brecha salarial entre trabajadores formales e informales en el Perú urbano, con una gran parte de esta brecha que no se puede atribuir a diferencias en las características observables de los trabajadores. También se observa una brecha salarial significativa entre trabajadores informales dentro y fuera del sector informal, siendo que los informales que trabajan en el sector formal ganan salarios notablemente mayores a los informales del sector informal (Rentería 2015).

Barco y Vargas (2010) buscan medir las diferencias de ingresos entre trabajadores formales e informales en el Perú, usando datos de la ENAHO del INEI para el periodo 2004-2012. La técnica utilizada para estimar las brechas salariales es la descomposición no paramétrica de Ñopo (2004), que divide la brecha en cuatro componentes: las diferencias en las características observables (ΔO), no observables (ΔU), retornos a características observables (ΔR) y distribución de características entre formales e informales dentro del soporte común (ΔX). Esta técnica es flexible al no asumir restricciones sobre el comportamiento de las variables aleatorias en el modelo. Los hallazgos muestran que, para los asalariados, el 55% de la brecha salarial se debe a factores dentro del soporte común, como edad, nivel educativo, estado civil y dominio geográfico. Para los independientes, estas diferencias explican el 37% y el 27% de la brecha total, respectivamente. Se observa que los asalariados formales tienen casi nula



probabilidad de pasar a empleos asalariados informales, aunque podrían elegir empleos informales dentro de empresas formales. En promedio, el componente no explicado de las brechas de ingresos entre formales e informales dentro de la misma categoría ocupacional es de un 35% (Barco y Vargas 2010).

2.2. MARCO TEÓRICO

2.2.1. Teoría Neoclásica del Mercado de trabajo

La informalidad laboral ha sido objeto de estudio desde los años setenta, cuando se abordó desde la perspectiva del capital humano de Becker (1975). Según esta perspectiva, la competencia en el mercado laboral determina que los ingresos reflejen los niveles de productividad, que a su vez distinguen entre empleos formales e informales. Por ello, es relevante analizar la teoría de la oferta y la demanda de trabajo, para comprender el origen de la informalidad laboral y la brecha salarial.

La relación entre informalidad y baja productividad se explicaría por la oferta y la demanda laboral. La oferta depende de factores como la experiencia y la educación de los trabajadores. El tamaño de las empresas influye en la demanda; las más pequeñas enfrentan obstáculos para obtener crédito y financiar su capital. La baja inversión por empleado en sus actividades económicas implica que el rendimiento del trabajo sea menor y, por lo tanto, que se reduzca la formalización del empleo.

La empresa tiene como objetivo obtener el mayor nivel de rentabilidad posible de la venta de productos $Q_x = f(K, L)$, cuya producción depende de los factores capital (K) y trabajo (L). Siguiendo el modelo de Nicholson (2008), se supone que la empresa actúa de forma racional y se enfrenta a un costo por el uso



de estos factores. Por simplicidad, solo se analizan el capital y el trabajo. Así, se formula la siguiente ecuación:

$$\text{Máx } Pf(k, L)$$

$$\text{s. a: } CT = wL + rK$$

La empresa busca maximizar su nivel de rentabilidad produciendo el bien X , cuyo precio P está dado por el mercado. Para ello, utiliza dos factores de producción: trabajo y capital, cuyos costos son w y r respectivamente. El costo total CT depende de la cantidad de trabajo y capital empleados. La función objetivo de la empresa es:

$$\mathcal{L} = Pf(k, L) - \lambda(CT - wL - rK)$$

Para hallar el máximo de la función objetivo, se debe derivar e igualar a cero. Esta es la condición necesaria de primer orden:

$$\frac{\partial \mathcal{L}}{\partial K} = f_k(k, L) - \lambda w = 0$$

$$\frac{\partial \mathcal{L}}{\partial L} = f_L(k, L) - \lambda r = 0$$

Teniendo en cuenta las ecuaciones anteriores:

$$TMS_{K,L} = \frac{PMg_L}{PMg_K} = \frac{f_L(k, L)}{f_k(k, L)} = \frac{w}{r}$$

Los factores productivos de capital y trabajo tienen distintas combinaciones óptimas que determinan sus salarios según su productividad marginal. Si se consideran dos tipos de empresa: una mediana y otra pequeña,

podemos hallar diferencias en la productividad marginal y los salarios de cada factor.

$$PMg * L = g\left(\frac{K}{L}\right) = g\left(\frac{w}{r}\right) \rightarrow \text{Empresa pequeña}$$

$$PMg * L = h\left(\frac{K}{L}\right) = h\left(\frac{w}{r}\right) \rightarrow \text{Empresa mediana}$$

$$g\left(\frac{w}{r}\right) < h\left(\frac{w}{r}\right)$$

Las empresas pequeñas enfrentan dificultades para acceder al mercado financiero y a la tecnología, lo que limita su capacidad de contratar a trabajadores productivos y ofrecer salarios competitivos $g\left(\frac{w}{r}\right)$. Aunque algunas puedan atraer a trabajadores con alta productividad, estos no recibirán una remuneración acorde. Por lo tanto, estas empresas tienen más probabilidades de operar en la informalidad. Por el contrario, las empresas medianas cuentan con más recursos financieros y tecnológicos, lo que les permite seleccionar a trabajadores con mayor productividad y pagarles salarios más altos $h\left(\frac{w}{r}\right)$. Esto genera una brecha de ingresos entre los trabajadores de ambos tipos de empresas, siendo los de la empresa pequeña la más vulnerables a ser informales en comparación con una empresa mediana.

Becker (1975) afirma que el nivel de educación (capital humano) influye en la productividad y la calidad de los trabajadores, y que esto se refleja en sus salarios (calificación) y en su situación laboral (formal o informal). Según la perspectiva neoclásica, la informalidad es un fenómeno transitorio que surge por desequilibrios momentáneos o por dificultades como el desempleo involuntario. Sin embargo, en el largo plazo, el mercado se ajustaría por sí mismo, gracias a la

racionalidad de los agentes que buscan maximizar su beneficio en un entorno de movilidad perfecta e información.

2.2.2. El Modelo Teórico Ocio Consumo

Siguiendo a Krugman y Wells (2009) y Nicholson (2008), este modelo asume que cada individuo actúa racionalmente, es decir, busca optimizar una función de utilidad que depende de dos variables: Ocio (Horas que no se dedican al trabajo) y Consumo (Horas que se dedican a trabajar). El individuo busca maximizar su utilidad $U(O, C)$, que depende de las horas de ocio (O) y de consumo (C). La utilidad es una función cuasicóncava, lo que implica que tiene un único máximo. El problema de optimización está sujeto a dos restricciones: el tiempo total disponible y el ingreso total generado.

$$\text{Max } U(C, L)$$

$$\text{S. a } C + wL = V + wT \text{ con } L \leq T$$

El individuo busca maximizar su utilidad U , que depende del consumo de bienes C y del tiempo de ocio L . Su ingreso total se compone del salario por hora w multiplicado por las horas de trabajo H , más el ingreso no laboral V . El individuo tiene una restricción de tiempo T , que es la suma de las horas de ocio L y las horas de trabajo H . Para resolver este problema, se utiliza el método de Lagrange, con λ y μ como los multiplicadores asociados a las restricciones presupuestaria y temporal, respectivamente.

$$\mathcal{L}(C, L, \lambda, \mu) = U(C, L) + \lambda[V + wT - (C + wL)] + \mu(T - L)$$



Una forma de hallar los valores óptimos de O y L es derivar la función objetivo con respecto a estas variables e igualar a cero. Así se obtienen las condiciones de primer orden:

$$U'C = \lambda$$

$$U'L = \lambda w + \mu$$

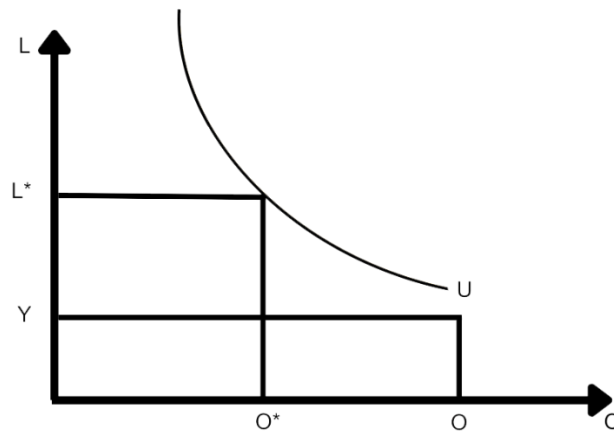
Una forma de explicar el punto óptimo es cuando el individuo iguala la tasa a la cual está dispuesto a intercambiar ocio por consumo con la tasa a la cual el mercado le permite hacer ese intercambio, en otras palabras, la condición para que el individuo ofrezca su trabajo es que el salario real sea superior o igual al valor de la TMS, que representa la cantidad de ocio que está dispuesto a sacrificar por una unidad adicional de consumo. Esto se cumple cuando:

$$\frac{U_{mgL}}{U_{mgC}} = w^*$$

La decisión de participar o no en el mercado laboral depende de la utilidad marginal del trabajo (U_{mgL}) y la utilidad marginal del consumo (U_{mgC}), que miden el beneficio adicional que se obtiene al trabajar y consumir una unidad más, respectivamente. El salario de reserva (w^*) es el nivel de ingreso que hace que el individuo sea indiferente entre trabajar o no. En el gráfico siguiente, se muestra que el individuo optimiza su elección de ocio (O^*) y consumo o horas de trabajo (L^*) cuando el salario de reserva es igual a la tasa marginal de sustitución (TMS) entre ocio y consumo. En ese punto, la curva de utilidad es tangente a la restricción presupuestaria.

Figura 1

Nivel de equilibrio



Fuente: Krugman y Wells (2009)

La variable que determina la decisión de trabajar de un individuo es el salario de reserva, es decir, el salario mínimo que aceptaría para ofrecer su trabajo. Si el salario de mercado o el salario que se le ofrece es igual o mayor al salario de reserva, el individuo decide trabajar; si no, prefiere no hacerlo. Esta es la lógica que subyace a esta variable.

- $Y_i = 1$ si $W_i > W_i^*$: Decide participar en el mercado laboral
- $Y_i = 0$ si $W_i < W_i^*$: Decide no participar en el mercado laboral

2.2.3. Teoría de segmentación, en relación a la brecha de ingreso

La teoría de segmentación laboral surgió como una respuesta a las limitaciones de la teoría neoclásica para explicar fenómenos socioeconómicos. Este enfoque sostiene que el mercado laboral no es homogéneo, sino que se divide en segmentos con características y condiciones distintas. Según Krugman y Wells (2009), los mercados segmentados son comunes y los trabajadores se clasifican según habilidades o experiencia.



La segmentación laboral se basa en factores como habilidades, tipo de trabajo, condiciones laborales y disponibilidad de trabajadores en cada segmento. Blanchard et al. (2012) añaden que esta segmentación puede surgir por falta de información, disparidades en formación y habilidades, y restricciones en la movilidad geográfica.

La teoría sostiene que la falta de oportunidades formales puede impulsar la participación en empleo informal, perpetuando así la segmentación. Krugman y Wells (2009) indican que, aunque el empleo informal es una alternativa para algunos, enfrenta salarios más bajos y condiciones precarias. La estructura del mercado laboral afecta la distribución de ingresos y la movilidad social. Blanchard et al. (2012) destacan que la educación es clave para reducir la informalidad, mejorando las oportunidades laborales.

Desde una perspectiva económica, Galiani y Weinschelbaum (2007) plantean un modelo con segmentos formales, informales superiores e informales inferiores, donde los individuos eligen empleos maximizando su utilidad. La existencia de barreras afecta la movilidad, generando segmentación y disparidades salariales. La teoría sugiere que beneficios pecuniarios y no pecuniarios, como el capital humano y la flexibilidad familiar, influyen en la elección entre empleo formal e informal.

Según Galiani y Weinschelbaum (2007), el mercado laboral se compone de tres segmentos: uno formal y dos informales, uno superior y otro inferior. El modelo asume que los individuos eligen el tipo de empleo (formal o informal) que les brinda mayor satisfacción. La satisfacción de elegir un empleo formal se expresa mediante la siguiente función de utilidad:



$$U_f = U(w_f L_f, B_f) = (w_f L_f + B_f)(1 - E)$$

El trabajo formal ofrece una utilidad de tres factores: el salario por una hora de trabajo (w_f), la cantidad de trabajo (L_f) y la rentabilidad no dineraria (B_f) que se obtienen en ese sector. Sin embargo, existen barreras de entrada (E) que restringe el acceso a los formales. Por el contrario, la utilidad del trabajo se define de esta manera:

$$U_i = U(w_i L_i, B_i) = (w_i L_i + B_i)$$

La utilidad de optar por un trabajo informal depende del salario, el esfuerzo y los beneficios no monetarios que se obtienen en ese sector, según la siguiente fórmula: $U_i = w_i L_i + B_i$. A partir de esta fórmula, Garavito (2011) plantea unas proposiciones que se basan en la maximización de la utilidad.

La Proposición 1 afirma que los mercados laborales se dividen en dos segmentos: el formal, que ofrece mejores condiciones, y el informal, que ofrece peores. Esto implica que los trabajadores informales desearían volverse formales, pero no pueden hacerlo porque hay restricciones para poder entrar.

La segunda proposición plantea que los mercados laborales se dividen en varios segmentos si existe un segmento informal de baja calidad con obstáculos para acceder y un segmento informal de alta calidad con trabajadores que optan por la informalidad de forma voluntaria.

La Proposición 3 señala que, al tener en cuenta las barreras de entrada, los beneficios monetarios y no monetarios influyen en la elección del tipo de empleo. La educación, entendida como inversión en capital humano, se relaciona



directamente con la disminución de la informalidad laboral (Blanchard et al. 2012).

Se puede concluir, que la teoría de segmentación de mercado laboral brinda un enfoque integral para comprender las diferencias de ingreso, integrando elementos estructurales y sociodemográficos que influyen en la movilidad y las oportunidades laborales.

2.3. MARCO CONCEPTUAL

Empleo: Para el INEI (2017), el empleo se refiere a la parte de la población en edad de trabajar que tiene una ocupación, ya sea en áreas urbanas o rurales. Dentro del empleo se distinguen dos grupos: la Población Económicamente Activa (PEA), que comprende a las personas que trabajan o buscan trabajo, y la Población Económicamente No Activa (No PEA), que incluye a las personas que no trabajan ni buscan trabajo por diversas razones, como ser estudiantes, amas de casa, pensionistas, jubilados, rentistas o discapacitados permanentes para trabajar, entre otras.

Empleo Informal: Según el INEI (2017), la define como aquel que no cuenta con los beneficios laborales que establece la ley, los cuales pueden variar según el país e incluir aspectos como seguridad social, vacaciones pagadas y licencia por enfermedad. Esta definición abarca varias categorías ocupacionales del empleo informal, como los empleadores, los trabajadores por cuenta propia y los trabajadores familiares no pagados que forman parte del sector informal, así como los asalariados que no cuentan con seguridad social financiada por su empleador, independientemente de si la unidad productiva es formal o informal. El INEI (2017) amplía esta definición para abarcar tanto el empleo en el sector



informal como el empleo informal en empresas formales o fuera del sector informal. En síntesis, se refiere a trabajadores que no gozan de beneficios laborales, como seguridad social y gratificaciones, y que laboran en empresas no registradas oficialmente.

Población Económicamente activa: la población económicamente activa (PEA) está formada por las personas mayores de 14 años que participan o buscan participar en el mercado laboral durante un período específico. Estas personas aportan o están dispuestas a aportar a la producción de bienes y servicios. La PEA incluye a los ocupados y a los desocupados que realizan una búsqueda activa de empleo (INEI 2017).

La Población Económicamente Inactiva (PEI): Según el INEI (2017), la PEI es la sigla que designa a la población económicamente inactiva, es decir, a aquellas personas que, estando en edad de trabajar, no participan en el mercado de trabajo. Esto implica que no realizan ni buscan ningún tipo de actividad económica remunerada. La PEI está compuesta por estudiantes, jubilados o pensionistas, rentistas, personas dedicadas al cuidado del hogar, ancianos, etc.

Población Económicamente Activa ocupada: Según el INEI (2017), se considera que una persona está ocupada si cumple con al menos uno de los siguientes requisitos: i). Tener 14 años o más y haber realizado alguna actividad económica en el período de referencia. ii). Ser trabajador dependiente con empleo fijo que no trabajó la semana anterior por razones como vacaciones, huelga, licencia médica, licencia maternal, etc., siempre que reciba remuneración. iii). Ser trabajador independiente que se ausentó temporalmente del trabajo en el período de referencia, pero que continuó con su empresa o negocio en funcionamiento.

iv). Haber participado en alguna actividad económica en el período de referencia, aunque sea por una hora, y recibir pago en dinero o en especie.

Brecha de Ingreso: Es la diferencia entre el ingreso que se obtiene en el empleo formal y el que se percibe en el informal.

2.4. MÉTODOS EMPÍRICOS PARA LA ESTIMACIÓN

2.4.1. Metodología econométrica Logit y Probit

Modelo Logit

En base a lo desarrollado por Gujarati y Porter (2010) y Wooldridge (2020), nos indica que es una forma de estudiar la asociación entre una variable binaria de respuesta y varias variables explicativas. El Modelo Logit se basa en una ecuación que expresa la relación entre estas variables:

$$P(Y = 1) = F(Z_i) = \frac{e^{Z_i}}{1 + e^{Z_i}} = F(X_i\beta') = \frac{e^{X_i\beta'}}{1 + e^{X_i\beta'}}$$

$$\text{Donde } Z_i = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k$$

En este modelo utiliza la función de distribución logística, denotada por el símbolo Λ , como la función de distribución acumulativa (FDA) que se emplea. El Modelo Logit establece una relación entre la variable dicotómica Y_i y las variables X_{2i} a X_{ki} mediante la ecuación:

$$Y_i = \frac{1}{1 + e^{-(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k)}} + u_i$$

Se asume que $E(u_i) = 0$. Esto implica que la variable dependiente es binaria, lo que se puede traducir de la siguiente manera:

$$P(Y_i = 1) = E(Y_i|X_i) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k)}} + u_i$$

Aquí, F es la función de distribución logística, β_i son los parámetros del modelo, X_i son las variables independientes y u_i es el término de error.

El objetivo del Modelo Logit es estimar la probabilidad P_i de que la variable Y_i tome el valor 1, en función de los valores de las variables X_i . De Esta forma:

$$Prob(Y_i = 1|X_i) = P_i$$

Efectos marginales del Modelo Logit

La ecuación de probabilidad nos permite hallar los efectos marginales de variables que pueden ser cualitativas o cuantitativas, una vez que hemos ajustado el modelo. Estos se obtienen de la siguiente manera:

$$\frac{\partial P(X_i \beta')}{\partial X_{ki}} = F(X_i \beta') \cdot [1 - F(X_i \beta')]$$

En función de los valores de X, estos valores cambian y se pueden examinar para distintos valores de X. Otra forma de expresarlos es mediante:

$$\frac{\partial P(Y = 1 | X)}{\partial X} = \beta_1 \cdot P(Y = 1 | X) \cdot [1 - P(Y = 1 | X)]$$

Una forma de medir el impacto de una variable categórica sobre la probabilidad dependiente es comparar las probabilidades cuando la variable explicativa cambia de 0 a 1.

El Modelo Probit

El Modelo Probit es una técnica estadística así como lo mencionan Gujarati y Porter (2010) y Wooldridge (2020), que utiliza la función de distribución normal acumulativa para modelar una variable de respuesta binaria. La ecuación del Modelo Probit es:

$$Y_i = \int_{-\infty}^{Z_i} \frac{1}{\sqrt{2\pi}} e^{-\frac{s^2}{2}} ds u_i$$

El modelo Probit se basa en el índice $Z_i = X_i\beta$, donde X_i son las variables explicativas, β_i son los parámetros. La probabilidad de que la variable dependiente sea 1 se obtiene integrando la función de densidad normal estándar desde $-\infty$ hasta Z_i . Se puede escribir el modelo como:

$$Y_i = \Phi(X_i\beta) + u_i = \Phi(Z_i) + u_i$$

$$\text{Donde } Z_i = \beta_0 + \beta_1x_1 + \dots + \beta_kx_k$$

Además, Φ es la función de distribución acumulativa normal estándar, β_i representando a los parámetros, X_i son las variables independientes y u_i es el término de error

En este modelo, se asigna una probabilidad P_i de que la variable Y_i sea 1, dado los valores de las características X_i . Así que

$$\text{Prob}(Y_i = 1|X_i) = P_i$$

Si las variables X_i tienen los mismos valores, entonces la variable Y_i tienen una probabilidad de $(1 - P_i)$ de ser 0, porque la probabilidad total es 1. El caso que tenemos:

$$\text{Prob}(Y_i = 0|X_i) = (1 - P_i)$$

Efectos Marginales del Modelo Probit

La ecuación de probabilidad nos permite obtener los efectos marginales para variables de tipo cualitativo y cuantitativo, una vez que hemos estimado el modelo Probit. Estos se obtienen como:

$$\frac{\partial \Phi(X_i \beta')}{\partial X_{ki}} = \Phi(X_i \beta') \cdot [1 - \Phi(X_i \beta')]$$

La relación entre estos valores y los valores de X depende de cada valor de X que se considere. Para una variable cualitativa, el efecto marginal se mide como la diferencia de probabilidades cuando la variable explicativa cambia de 0 a 1.

2.4.2. Metodología para la estimación de la Brecha de Ingresos

Metodología de Oaxaca-Blinder para la obtención de brechas salariales entre el empleo formal e informal

La descomposición de Oaxaca-Blinder es una técnica estadística propuesta por Oaxaca (1973) y Blinder (1973). Este método se utiliza para explicar las diferencias en las medias de una variable dependiente (como los salarios) entre dos grupos (como el empleo formal e informal).

Para entender esto, comencemos con las ecuaciones de renta laboral para el trabajo informal y formal, basadas en las ecuaciones de Mincer (1974):

Ecuación de ingreso laboral para el empleo formal:

$$Y_i^F = \beta_0^F + \sum_{j=1}^F \beta_j^F \beta_{ji}^F \varepsilon_0^F$$

Ecuación de ingreso laboral para el empleo informal:

$$Y_i^I = \beta_0^I + \sum_{j=1}^I \beta_j^I \beta_{ji}^I \varepsilon_0^I$$

Además, podemos plantear la diferencia de ingresos siguiendo a Arpi y Arpi (2010), de la manera siguiente:

$$DI_{it} = \left(1 - \frac{\text{Ingreso laboral del empleo formal}}{\text{Ingreso laboral del empleo informal}} \right) * 100$$

Seguidamente, planteamos la diferencia salarial entre el trabajo informal y formal, que se obtiene de la resta de las ecuaciones de ingreso laboral del empleo formal e informal:

$$Y_i^F - Y_i^I = \beta_0^F - \beta_0^I + \sum_{j=1}^F \beta_j^F X_{ji}^F - \sum_{j=1}^I \beta_j^I X_{ji}^I$$

Esta fórmula muestra cómo se puede expresar la variación en términos de variables explicativas (X) y los parámetros estimados (β), de la siguiente manera:

$$\begin{aligned} Y_i^F - Y_i^I &= \beta_0^F - \beta_0^I + \sum_{j=1}^F \beta_j^F (X_{ji}^F - X_{ji}^I) + \sum_{j=1}^I X_{ji}^I (\beta_j^F - \beta_j^I) \\ &+ \left(\sum_{j=1}^F \beta_j^F - \sum_{j=1}^I \beta_j^I \right) * (\beta_j^F - \beta_j^I) \end{aligned}$$

Así, esta fórmula muestra que tiene 4 partes:

La primera: $(\beta_0^F - \beta_0^I)$: Esta es la diferencia en términos independientes, que hace referencia a la presencia de variables ausentes en el modelo

La segunda: $\sum_{j=1}^F \beta_j^F (X_{ji}^F - X_{ji}^I)$: Esta diferencia representa a los beneficios particulares del trabajo formal.

La tercera: $\sum_{j=1}^I X_{ji}^I (\beta_j^F - \beta_j^I)$: Esta diferencia representa a los beneficios particulares del trabajo informal.

La cuarta: $(\sum_{j=1}^F \beta_j^F - \sum_{j=1}^I \beta_j^I) * (\beta_j^F - \beta_j^I)$: Este término está asociado al hecho de que las diferencias en dotaciones y coeficientes existen simultáneamente entre los dos grupos.

Que también se podría expresar de la siguiente manera, siguiendo a Esparta y Rivera (2020), en términos de componentes observables y componentes no observables:

$$\Delta \overline{LnW}_{F-I} = \underbrace{\hat{\beta}_I (\bar{X}_F - \bar{X}_I)}_{\Delta O} + \underbrace{\hat{X}_F (\hat{\beta}_F - \hat{\beta}_I)}_{\Delta NO}$$

Donde:

$\hat{\beta}$: simboliza a los estimadores de los ingresos y \bar{X} simboliza el promedio de cada una de las variables explicativas de trabajadores formales (F) e informales (I), respectivamente.

ΔO : Mide el diferencial salarial atribuido a las diferencias en las características observadas entre ambos grupos.

ΔNO : Mide el diferencial salarial atribuido a las diferencias a las características no observadas, conocido en el mundo de las investigaciones como discriminación salarial.

Metodología de Ñopo para la obtención de brechas salariales entre el empleo formal e informal

La metodología desarrollada por Ñopo (2004), ofrece una forma flexible de construir la distribución salarial, evitando asumir restricciones sobre el comportamiento de variables aleatorias en el modelo. La metodología permite descomponer la brecha salarial entre trabajadores formales e informales con atributos similares, y así identificar componentes específicos de la diferencia de ingresos, lo cual aporta más información que solo observar la brecha media.

Iniciamos señalando que $F^I(\cdot)$ y $F^F(\cdot)$ representan las funciones de distribución acumulativa condicionadas de las características individuales X , según si se trata de informal o formal, respectivamente. Para el empleo formal, tenemos:

$$u^F(S) = \int_S dF^F(x)$$

De manera análoga, para el empleo informal, tenemos:

$$u^I(S) = \int_S dF^I(x)$$

La conexión entre estas variables es descrita por las funciones $g^F(\cdot)$ y $g^I(\cdot)$, que representan el ingreso esperado condicionado a las características y al estatus de formalidad, de la siguiente manera:

$$E[Y|F, X] = g^F(x) \rightarrow E[Y|F] = \int_{S^F} g^F(x) dF^F(x)$$

y

$$E[Y|I, X] = g^I(X) \rightarrow E[Y|I] = \int_{S^I} g^I(x) dF^I(x)$$

Donde, S^F representa el rango de la distribución de las características para los trabajadores formales, y S^I representa el rango de la distribución de las características para los trabajadores informales. Así, la brecha de ingresos se expresaría de la siguiente manera:

$$\Delta \equiv E[Y|F] - E[Y|I]$$

Expresada de otra forma:

$$\Delta \equiv \int_{S^F} g^F(x) dF^F(x) - \int_{S^I} g^I(x) dF^I(x)$$

Considerando que el rango de distribución de características para trabajadores formales S^F , difiere del rango de distribución de las características para los trabajadores informales S^I , cada integral se divide en dos componentes: uno en la intersección de los rangos y otro fuera del rango compartido, de la manera siguiente:

$$\Delta \equiv \left[\int_{S^F \cap S^I} g^I(x) dF^I(x) + \int_{S^F \setminus S^I} g^I(x) dF^I(x) \right] - \left[\int_{S^F \cap S^I} g^F(x) dF^F(x) - \int_{S^F \setminus S^I} g^F(x) dF^F(x) \right]$$

La brecha salarial relativa (Δ_{Total}) del ingreso laboral promedio entre formales e informales se encuentra conformada por cuatro componentes, luego de algunos artificios matemático se puede obtener lo siguiente:

$$\Delta_{Total} = \Delta_{formal} + \Delta_x + \Delta_d + \Delta_{Informal}$$



La explicación de cada uno de los elementos es la siguiente:

Δ_{formal} : Brecha explicada por las diferencias entre grupos de trabajadores formales con y sin contrapartes similares en empleo informal, identificando aquellos fuera del soporte común sin contrapartes informales.

Δ_x : Diferencia atribuida a las disparidades en la distribución de características observables entre trabajadores formales e informales que comparten un conjunto común de características.

Δ_d : Parte "no explicada" de la brecha salarial entre formales e informales dentro del soporte común, atribuida a discriminación y características no observables que influyen en los ingresos.

$\Delta_{Informal}$: Parte de la brecha explicada por diferencias entre dos grupos de trabajadores informales, análogo a Δ_{formal} . Se refiere a la diferencia entre trabajadores informales con características emparejables a formales y aquellos sin contrapartes similares, revelando la porción que desaparecería sin informales sin equivalentes formales.

2.5. HIPÓTESIS DE LA INVESTIGACIÓN

2.5.1 Hipótesis general

La variable educación y la creación de nuevas empresas influye sobre el empleo informal y en la diferencia de ingresos entre ocupaciones informales y formales en Perú, periodo 2022.

2.5.2. Hipótesis específicos



- Existe una relación significativa y negativa entre la ocupación informal y las variable educación y creación de nuevas empresas en Perú, periodo 2022.
- Las variables educación y creación de nuevas empresas explican significativamente la diferencia de ingresos entre ocupaciones informales y formales en Perú, periodo 2022.

CAPÍTULO III

MATERIALES Y MÉTODOS

3.1. UBICACIÓN GEOGRÁFICA DEL ESTUDIO

El Perú como lugar de estudio, la ubicación geográfica del Perú en la región noroeste de América del Sur le confiere una diversidad económica, social y cultural única.

Figura 2

Mapa del Perú



Fuente: página web: [https:// viajaporperu.com/](https://viajaporperu.com/)



3.2. MÉTODO Y TIPO DE INVESTIGACIÓN

En base a lo abordado por Mendoza (2014) y Hernández et al. (2023), para investigar el fenómeno de la informalidad, se propone una hipótesis que se ajuste a la teoría económica. Se utilizará un método hipotético deductivo, que infiere de lo general a lo particular, y un enfoque descriptivo, que caracteriza la informalidad según variables como la educación y la creación de nuevas empresas. El método es no experimental, ya que no se pueden controlar las variables, y transversal, ya que se basa en datos de un momento determinado.

3.3. PROCEDENCIA DEL MATERIAL UTILIZADO

La fuente de información para este estudio es la Encuesta Nacional de Hogares que realizó el Instituto Nacional de Estadística e Informática (INEI) para el año 2022, tomando como referencia el módulo 1 características de vivienda y del hogar, modulo 2 características de los miembros del hogar, el módulo 5 empleo e ingreso y el módulo 34 Sumarias.

3.4. POBLACIÓN Y MUESTRA DE ESTUDIO

La población de estudio, según datos del INEI (2022), estará conformada por individuos en edad de trabajar (PET) que abarcan desde los 14 hasta los 65 años de edad, específicamente aquellos que forman parte de la población económicamente activa ocupada (PEA - ocupada). La fuente de información provendrá de la Encuesta Nacional de Hogares (ENAH). Según la ficha técnica proporcionada por el INEI (2022), se establece una muestra de 36,848 viviendas particulares para el año 2022. Esta muestra es de tipo probabilística, estratificada, de áreas, multietápica y única para cada departamento objeto de estudio. Es importante señalar que la muestra se redujo a 54,582.00

observaciones, ya que se consideró únicamente a los trabajadores en edad de trabajar, y se tuvo en cuenta la presencia de valores faltantes (missing) en la muestra.

3.5. PROCEDIMIENTO

3.5.1. METODOLOGÍA PARA EL OBJETIVO 1

El modelo econométrico Logit y Probit

Se emplea la metodología Logit y Probit en esta investigación debido a la amplia referencia de estudios previos que han utilizado estos modelos. La elección de estos modelos se justifica por su capacidad para modelar relaciones probabilísticas, lo que resulta adecuado para nuestro objetivo de analizar la relación entre la variable "Empleo informal" y las variables "Educación" y "Creación de nuevas empresas". Siguiendo la metodología propuesta por Gujarati y Porter (2010), así como las pautas establecidas por Calatayud (2019), el enfoque metodológico se estructura de la siguiente manera:

Logit

$$\begin{aligned} PLB_i &= \ln\left(\frac{P_i}{1 - P_i}\right) = \beta_0 + \beta_1 X_1 + \dots + \beta_K X_K \\ PLB_i &= \ln\left(\frac{P_i}{1 - P_i}\right) \\ &= \beta_0 + \beta_1 (EDU) + \beta_2 (EXP) + \beta_3 (EXP^2) + \beta_4 (SEXO) \\ &\quad + \beta_5 (NEMP) + \beta_6 (TEMP) + \varepsilon_i \end{aligned}$$

Probit

$$Prob(PLB = 1 | X) = \Phi(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k)$$



$$\begin{aligned} Prob(PLB = 1 | X) \\ = \Phi(\beta_0 + \beta_1 (EDU) + \beta_2 (EXP) + \beta_3 (EXP^2) + \beta_4 (SEXO) \\ + \beta_5 (NEMP) + \beta_6 (TEMP) + \varepsilon_i) \end{aligned}$$

PLB = Variable independiente: (1 = informal, 0 = formal)

X_i = Variables independientes o regresores

ε_i = Término de perturbación estocástica

β_i = Representa los parámetros o regresores del modelo, donde:

EDU: Años de educación alcanzados

EXP: Experiencia potencial

*EXP*²: Experiencia potencial al cuadrado

SEXO: Sexo

NEMP: Número de nuevas empresas

TEMP: Número de trabajadores en la empresa

De acuerdo a Calatayud (2019), la experiencia potencial puede calcularse a partir de la edad y la educación siguiendo la siguiente forma:

$$EXP_i = EDAD_i - EDU_i - 4$$

Para la interpretación de resultados, se interpretó los efectos marginales que se obtienen a través de la derivada parcial de cada variable.

3.5.2. Metodología para el objetivo 2

Técnica de descomposición de Oaxaca – Blinder para la obtención de brechas salariales entre el empleo formal e informal

De acuerdo a Calatayud (2019), para poder hallar la brecha salariales entre los trabajadores formales e informales, se requiere comprobar la siguiente hipótesis nula:

$$H_0 = Y_i^{formal} - Y_i^{informal} = 0$$

Teniendo en cuenta la hipótesis alterna

$$H_a = Y_i^{formal} - Y_i^{informal} \neq 0$$

Así, planteando la regresión con la muestra completa:

$$\ln(Y_i) = \beta_0 + \beta_1 (EDU) + \beta_2 (EXP) + \beta_3 (EXP^2) + \beta_4 (SEXO) \\ + \beta_5 (NEMP) + \beta_6 (TEMP) + \varepsilon_i$$

Posteriormente, se llevarán a cabo estimaciones por separado para el grupo de empleo formal y el empleo informal, de la siguiente manera:

$$\ln(Y_i^{FORMAL}) = \beta_0^{for} + \beta_1^{for} (EDU_i^{for}) + \beta_2^{for} (EXP_i^{for}) + \beta_3^{for} (EXP_i^{for})^2 \\ + \beta_4^{for} (SEXO_i^{for}) + \beta_5^{for} (NEMP_i^{for}) + \beta_6^{for} (TEMP_i^{for}) \\ + \varepsilon_i^{for}$$

$$\ln(Y_i^{INFORMAL}) \\ = \beta_0^{inf} + \beta_1^{inf} (EDU_i^{inf}) + \beta_2^{inf} (EXP_i^{inf}) + \beta_3^{inf} (EXP_i^{inf})^2 \\ + \beta_4^{inf} (SEXO_i^{inf}) + \beta_5^{inf} (NEMP_i^{inf}) + \beta_6^{inf} (TEMP_i^{inf}) \\ + \varepsilon_i^{inf}$$

Además, en el análisis de esta técnica implica desagregar en dos componentes, la primera en características no observables y la segunda en características observables. Es decir.

$$CO = [\beta_0^{for} \beta_1^{for} \beta_2^{for} \beta_3^{for}] \left[\begin{array}{c} 1 \\ \overline{EDU^{for}} \\ \overline{EXP^{for}} \\ \overline{(EXP^2)^{for}} \end{array} \right] - \left[\begin{array}{c} 1 \\ \overline{EDU^{inf}} \\ \overline{EXP^{inf}} \\ \overline{(EXP^2)^{inf}} \end{array} \right]$$
$$CNO = [1 \overline{EDU^{inf}} \overline{EXP^{inf}} \overline{(EXP^2)^{inf}}] \left[\begin{array}{c} \beta_0^{for} \\ \beta_1^{for} \\ \beta_2^{for} \\ \beta_3^{for} \end{array} \right] - \left[\begin{array}{c} \beta_0^{inf} \\ \beta_1^{inf} \\ \beta_2^{inf} \\ \beta_3^{inf} \end{array} \right]$$

Donde: $\overline{EDU^{for}}$ y $\overline{EDU^{inf}}$ = Promedio de los años de educación del grupo de formales e informales y de manera análoga para los demás elementos, que representan, los parámetros y los promedios.

Metodología de Ñopo para la obtención de brechas salariales entre el empleo formal e informal

Para poder hallar la brecha salarial entre los trabajadores formales e informales, se requiere comprobar la siguiente hipótesis nula:

$$H_0 = u_i^{formal} - u_i^{informal} = 0$$

Teniendo en cuenta la hipótesis alterna

$$H_a = u_i^{formal} - u_i^{informal} \neq 0$$

Así, planteando la regresión con la muestra completa: Para el empleo formal, tenemos:

$$u^F(S) = \int_S dF^F(\beta_0 + \beta_1(EDU) + \beta_2(EXP) + \beta_3(EXP^2) + \beta_4(SEXO) \\ + \beta_5(NEMP) + \beta_6(TEMP) + \varepsilon_i^{for})$$

De manera análoga, para el empleo informal, tenemos:

$$u^I(S) = \int_S dF^I(\beta_0 + \beta_1(EDU) + \beta_2(EXP) + \beta_3(EXP^2) + \beta_4(SEXO) \\ + \beta_5(NEMP) + \beta_6(TEMP) + \varepsilon_i^{infor})$$

La conexión entre estas variables es descrita por las funciones $g^F(\cdot)$ y $g^I(\cdot)$, que representan el ingreso esperado condicionado a las características y al estatus de formalidad, de la siguiente manera:

$$X = \beta_0 + \beta_1(EDU) + \beta_2(EXP) + \beta_3(EXP^2) + \beta_4(SEXO) + \beta_5(NEMP) \\ + \beta_6(TEMP) + \varepsilon_i^{infor}$$

$$E[Y|F, X] = g^F(X) \rightarrow E[Y|F] = \int_{S^F} g^F(X) dF^F(X)$$

y

$$E[Y|I, X] = g^I(X) \rightarrow E[Y|I] = \int_{S^I} g^I(X) dF^I(X)$$

Donde, S^F representa el rango de la distribución de las características para los trabajadores formales, y S^I representa el rango de la distribución de las características para los trabajadores informales. Así, la brecha de ingresos se expresaría de la siguiente manera:

$$\Delta \equiv E[Y|F] - E[Y|I]$$

Expresada de otra forma:

$$\Delta \equiv \int_{S^F} g^F(x) dF^F(x) - \int_{S^I} g^I(x) dF^I(x)$$

Considerando que el rango de distribución de características para trabajadores formales S^F , difiere del rango de distribución de las características para los trabajadores informales S^I , cada integral se divide en dos componentes: uno en la intersección de los rangos y otro fuera del rango compartido, de la manera siguiente:

$$\Delta \equiv \left[\int_{S^F \cap S^I} g^I(x) dF^I(x) + \int_{S^F \cap S^I} g^I(x) dF^I(x) \right] - \left[\int_{S^F \cap S^I} g^F(x) dF^F(x) - \int_{S^F \cap S^I} g^F(x) dF^F(x) \right]$$

La brecha salarial relativa (Δ_{Total}) del ingreso laboral promedio entre formales e informales se encuentra conformada por cuatro componentes, luego de algunos artificios matemático se puede obtener lo siguiente:

$$\Delta_{Total} = \Delta_{formal} + \Delta_x + \Delta_d + \Delta_{Informal}$$

La explicación de cada uno de los elementos es la siguiente:

Δ_{formal} : Brecha explicada por las diferencias entre grupos de trabajadores formales con y sin contrapartes similares en empleo informal, identificando aquellos fuera del soporte común sin contrapartes informales.

Δ_x : Diferencia atribuida a las disparidades en la distribución de características observables entre trabajadores formales e informales que comparten un conjunto común de características.

Δ_d : Parte "no explicada" de la brecha salarial entre formales e informales dentro del soporte común, atribuida a discriminación y características no observables que influyen en los ingresos.

$\Delta_{informal}$: Parte de la brecha explicada por diferencias entre dos grupos de trabajadores informales, análogo a Δ_{formal} . Se refiere a la diferencia entre trabajadores informales con características emparejables a formales y aquellos sin contrapartes similares, revelando la porción que desaparecería sin informales sin equivalentes formales

3.6. VARIABLES

Tabla 1.

Organización de variables

Variable	Descripción	Detalle
ingm	Ingreso laboral mensual (en logaritmos)	Soles
Ingh	Ingreso laboral por hora (en logaritmos)	Soles
ocupinf	Situación de informalidad (Ocup. Principal)	1. Empleo Informal 2. Empleo formal
edu	Años de educación alcanzados	Numero
exp	Experiencia potencial	Número
exp^2	Experiencia potencial al cuadrado	Número
sexo	Sexo	1. Hombre 2. Mujer
nemp	Empresas nuevas	Número
temp	Número de trabajadores en la empresa	1 = Hasta 20 2 = De 21 a 50 3 = De 51 a 100 4 = De 101 a 500

Fuente: Elaboración propia en base a ENAHO – INEI, 2022



CAPÍTULO IV

RESULTADOS Y DISCUSIÓN

4.1. RESULTADOS

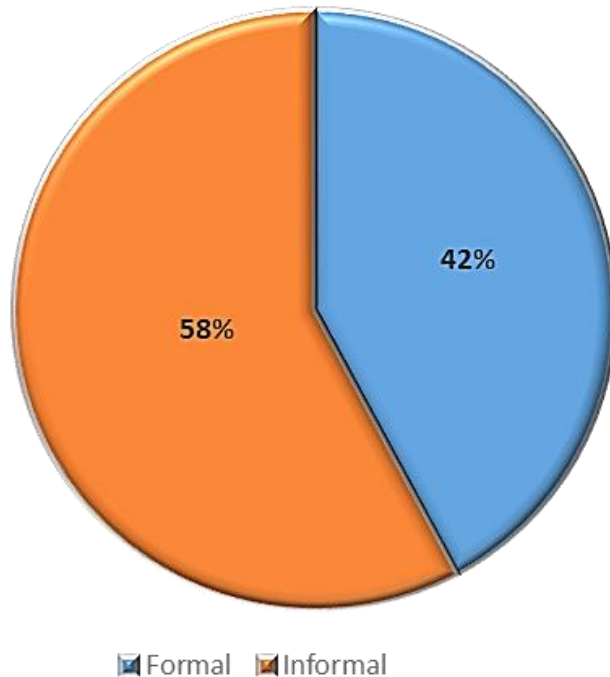
Este apartado presenta el análisis descriptivo de los resultados, enfocado en las variables de estudio: la educación y la creación de nuevas empresas. Para ello, se describe primero la relación entre los trabajadores formales e informales, y luego se examinan aspectos relacionados con la brecha de ingresos promedio por departamento, regiones naturales, ámbito, sexo y nivel educativo. Asimismo, se analiza la variación en la cantidad de trabajadores informales respecto a los formales, también por regiones naturales, ámbito, sexo y nivel educativo. Posteriormente, se plantea un modelo de regresión logística que permite responder al primer objetivo de la investigación. Finalmente, se realiza un análisis de brecha de ingreso utilizando las técnicas de Oaxaca – Blinder y la técnica de Ñopo, con las cuales se responde al segundo objetivo de la investigación

4.1.1. Análisis descriptivo de las variables

En la figura 3 se muestra el porcentaje de la población económicamente activa ocupada en un rango de edad de 14 a 65 años que participó en la muestra. Según la figura, el sector informal representa el 58% de la muestra, mientras que el sector formal representa el 42%.

Figura 3

Porcentaje de formalidad e informalidad



Fuente: elaboración propia en base a ENAHO – INEI, 2022

Según la figura 4, a nivel nacional, el departamento de Huánuco presenta la mayor brecha entre trabajadores formales e informales, pues el ingreso promedio de un formal supera en un 69% al de un informal. Por el contrario, el departamento con la menor brecha es Ica, con un 37%. En cuanto a la cantidad de trabajadores, la brecha más amplia se da en el departamento de Puno, donde el número de informales excede en un 70% al de formales. La brecha más estrecha se observa en el departamento de Moquegua, donde hay un 18% más de formales que de informales. Por último, respecto a la variación neta de creación de nuevas empresas, el departamento de Junín lidera con 4388 nuevas empresas, mientras que el departamento de Arequipa registra una disminución de 4860 empresas. Cabe destacar que el promedio de trabajadores por empresa oscila entre uno y dos.

Análisis descriptivo del Ingreso Promedio Número de casos por Regiones

Naturales

Tabla 2

Brecha del Ingreso Promedio y Número de Casos por Regiones Naturales

Depart.	Costa	Sierra	Selva	Total	Costa	Sierra	Selva	Total
Amazonas		57%	60%	59%		36%	63%	54%
Ancash	56%	65%		60%	20%	47%		33%
Apurímac		66%		66%		54%		54%
Arequipa	44%	56%		56%	62%			19%
Ayacucho		64%	67%	64%		43%	90%	48%
Cajamarca		65%	63%	64%		26%	41%	31%
Callao	43%			43%	-8%			-8%
Cusco		56%	33%	54%		36%	60%	39%
Huancavelica		63%		63%		55%		55%
Huánuco		69%	67%	69%		68%	24%	44%
Ica	37%			37%	-6%	0%		-6%
Junín		55%	55%	56%		29%	67%	38%
La Libertad	46%	53%		47%	9%	48%		16%
Lambayeque	50%			50%	27%			28%
Lima	46%	51%		46%	9%	68%		12%
Loreto			57%	57%			17%	17%
Madre de Dios			46%	46%			54%	54%
Moquegua	54%	66%		56%	-21%	3%		-18%
Pasco		57%	56%	57%		-1%	61%	18%
Piura	47%	70%		49%	29%	93%		33%
Puno		60%		61%		69%		70%
San Martín			57%	57%			43%	43%
Tacna	56%	62%		57%	40%	41%		40%
Tumbes	52%			52%	19%			19%
Ucayali			45%	45%			47%	47%
Nacional	46%	60%	56%	51%	14%	39%	45%	27%

Fuente: elaboración propia en base a ENAHO – INEI, 2022



En la tabla 2 el nivel de brecha entre trabajadores formales e informales es mayor en la sierra, donde el ingreso promedio de un formal supera en un 60% al de un informal. Le siguen la selva, con un 56% de diferencia, y la costa, con un 46%. Dentro de la costa, los departamentos con mayor brecha son Ancash y Tacna, ambos con un 56%, mientras que el de menor brecha es Ica con un 37%. En la sierra, la mayor brecha se da en Piura, con un 70%, y la menor en Lima, con un 51%. En la selva, la brecha más alta corresponde a Huánuco, con un 69%, y la más baja a Ica con un 37%.

Los datos de la tabla 2, del lado derecho muestran la diferencia porcentual entre los trabajadores informales y los formales en las tres regiones naturales del Perú: Costa, Sierra y Selva. Esta diferencia, o brecha, refleja el grado de informalidad laboral en cada región. De acuerdo con la tabla, la Selva presenta la mayor brecha con un 45%, seguida por la Sierra con un 39% y la Costa con un 14%. Sin embargo, dentro de cada región hay variaciones significativas entre los departamentos. Por ejemplo, en la Costa, Arequipa tiene la mayor brecha, llegando a un 62%, mientras que Moquegua tiene más trabajadores formales que informales, superándolos en 21%. En la Sierra, Piura tiene la mayor brecha, alcanzando un 93%, mientras que Pasco tiene un 1% más de trabajadores formales que informales. En la Selva, Ayacucho tiene la mayor brecha, con un 90%, y Loreto la menor, con un 17%.

Análisis descriptivo del Ingreso y Número de casos por Ámbito Rural y

Urbano

Tabla 3

Ingreso Promedio y Número de Casos por Ámbito

Depart.	Rural	Urbano	Total	Rural	Urbano	Total
Amazonas	67%	55%	59%	75%	32%	54%
Ancash	67%	57%	60%	74%	15%	33%
Apurímac	64%	61%	66%	78%	22%	54%
Arequipa	42%	56%	56%	73%	7%	19%
Ayacucho	65%	63%	64%	83%	25%	48%
Cajamarca	62%	61%	64%	77%	-16%	31%
Callao		43%	43%		-8%	-8%
Cusco	59%	52%	54%	78%	14%	39%
Huancavelica	63%	60%	63%	73%	16%	55%
Huánuco	72%	64%	69%	79%	13%	44%
Ica	28%	37%	37%	40%	-14%	-6%
Junín	59%	53%	56%	82%	14%	38%
La Libertad	47%	47%	47%	74%	6%	16%
Lambayeque	40%	51%	50%	57%	20%	28%
Lima	48%	45%	46%	77%	4%	12%
Loreto	65%	55%	57%	78%	-1%	17%
Madre de Dios	49%	45%	46%	79%	44%	54%
Moquegua	55%	55%	56%	40%	-27%	-18%
Pasco	59%	55%	57%	69%	-10%	18%
Piura	50%	47%	49%	75%	24%	33%
Puno	55%	62%	61%	89%	57%	70%
San Martín	64%	55%	57%	89%	25%	43%
Tacna	50%	58%	57%	40%	40%	40%
Tumbes	36%	52%	52%	63%	11%	19%
Ucayali	67%	43%	45%	88%	38%	47%
Nacional	58%	49%	51%	74%	12%	27%

Fuente: elaboración propia en base a ENAHO – INEI, 2022

En la tabla 3 el nivel de brecha entre trabajadores formales e informales es mayor en el ámbito Rural, donde el ingreso promedio de un formal supera en un

58% al de un informal. Le sigue el ámbito Urbano, con un 49% de diferencia. Dentro del ámbito Rural, el departamento con mayor brecha es Huánuco, con un 72%, mientras que el de menor brecha es Ica con un 28%. En el ámbito Urbano, la mayor brecha se da en Huánuco, con un 64% %, y la menor en Ica, con un 37%.

Los datos de la tabla 3, del lado derecho muestran la diferencia porcentual entre los trabajadores informales y los formales en los ámbitos Rural y Urbano. Esta diferencia, o brecha, refleja el grado de informalidad laboral en cada ámbito. De acuerdo con la tabla, el ámbito Rural presenta la mayor brecha con un 74%, seguida por el ámbito Urbano con un 12%. Sin embargo, dentro de cada ámbito hay variaciones. Por ejemplo, en el ámbito Rural, el departamento de Puno y San Martín tienen la mayor brecha, llegando ambos a un 89%, mientras que el departamento con menor brecha es Moquegua, alcanzando un 40%. En el ámbito Urbano, Puno tiene la mayor brecha, alcanzando un 57%, mientras que el departamento de Moquegua tiene más trabajadores formales que informales, superándolos en 27%.

Análisis descriptivo del Ingreso promedio y Número de casos por Sexo

Tabla 4

Ingreso Promedio y Número de Casos por Sexo

Depart.	Hombre	Mujer	Total	Hombre	Mujer	Total
Amazonas	57%	67%	59%	70%	24%	54%
Ancash	58%	64%	60%	37%	27%	33%
Apurímac	65%	68%	66%	57%	47%	54%
Arequipa	53%	58%	56%	18%	20%	19%
Ayacucho	58%	75%	64%	54%	39%	48%
Cajamarca	63%	68%	64%	37%	21%	31%



Depart.	Hombre	Mujer	Total	Hombre	Mujer	Total
Callao	36%	50%	43%	-24%	9%	-8%
Cusco	53%	59%	54%	50%	23%	39%
Huancavelica	57%	73%	63%	60%	47%	55%
Huánuco	66%	74%	69%	54%	28%	44%
Ica	31%	46%	37%	-7%	-5%	-6%
Junín	49%	66%	56%	38%	39%	38%
La Libertad	40%	57%	47%	15%	17%	16%
Lambayeque	41%	62%	50%	28%	29%	28%
Lima	41%	52%	46%	9%	16%	12%
Loreto	50%	68%	57%	29%	-2%	17%
Madre de Dios	42%	60%	46%	62%	38%	54%
Moquegua	52%	61%	56%	-26%	-9%	-18%
Pasco	50%	67%	57%	4%	34%	18%
Piura	49%	50%	49%	41%	20%	33%
Puno	55%	73%	61%	75%	59%	70%
San Martín	56%	63%	57%	52%	28%	43%
Tacna	53%	62%	57%	41%	39%	40%
Tumbes	43%	64%	52%	22%	15%	19%
Ucayali	43%	52%	45%	51%	39%	47%
Total	47%	58%	51%	30%	23%	27%

Fuente: elaboración propia en base a ENAHO – INEI, 2022

En la tabla 4 el nivel de brecha entre trabajadores formales e informales es mayor en las mujeres que en los Hombres, donde el ingreso promedio de un formal supera en un 58% al de un informal. En los Hombres esta cifra alcanza un 47% de diferencia. Además, dentro del grupo de los Hombres, el departamento con mayor brecha es Huánuco, con un 66%, mientras que el de menor brecha es Ica con un 31%. Dentro del grupo de las Mujeres, la mayor brecha se da en Ayacucho, con un 75%, y la menor en Ica, con un 46%.



Los datos de la tabla 4, del lado derecho muestran la diferencia porcentual entre los trabajadores informales y los formales para los grupos Hombres y Mujeres. Esta diferencia, o brecha, refleja el grado de informalidad laboral en cada ámbito. De acuerdo con la tabla, el grupo de los Hombres presenta la mayor brecha con un 30%, seguida por el grupo de las Mujeres con un 23%. Sin embargo, dentro de cada grupo hay variaciones. Por ejemplo, en el grupo de los Hombres, el departamento de Puno tiene la mayor brecha, llegando a un 75%, mientras que el departamento con menor brecha es Moquegua, alcanzando un -26%, lo que significa que Moquegua tiene más trabajadores formales hombres que informales. En el grupo de las Mujeres, Puno tiene la mayor brecha, alcanzando un 59%, mientras que el departamento de Moquegua tiene más trabajadores formales que informales, superándolos en 9%.

Análisis del Ingreso Promedio para el nivel educativo

Tabla 5

Ingreso Promedio por Nivel Educativo

Depart.	Secund. Incomp.	Secund. Compl.	Superior no universitaria Incompleta	Superior no universitaria completa	Sup. Univer. Incompl.	Sup. Univers. Comple.	Maestría Doctor.
Amazonas	21%	36%	54%	39%	65%	56%	32%
Ancash	55%	51%	48%	56%	32%	47%	25%
Apurímac	69%	50%	56%	50%	61%	41%	
Arequipa	16%	43%	50%	47%	53%	50%	67%
Ayacucho	42%	42%	48%	57%	66%	51%	
Cajamarca	38%	51%	80%	53%	40%	54%	
Callao	20%	24%	35%	46%	48%	44%	53%
Cusco	34%	43%	46%	52%	47%	47%	
Huancavelica	53%	53%	38%	52%	71%	2%	69%
Huánuco	70%	65%	24%	63%	15%	37%	33%
Ica	32%	27%	33%	26%	43%	40%	80%
Junín	37%	46%	40%	38%	51%	55%	15%
La Libertad	34%	38%	45%	40%	36%	29%	
Lambayeque	41%	32%	31%	44%	46%	40%	32%
Lima	33%	21%	24%	32%	35%	40%	26%
Loreto	26%	51%	36%	37%	60%	32%	14%
Madre de Dios	24%	34%	10%	46%	49%	63%	
Moquegua	32%	56%	50%	61%	69%	42%	21%
Pasco	69%	54%	54%	51%	50%	40%	
Piura	45%	39%	34%	44%	40%	16%	
Puno	77%	53%	33%	56%	41%	48%	61%
San Martín	50%	37%	32%	49%	51%	47%	61%
Tacna	44%	46%	36%	56%	68%	49%	
Tumbes	25%	39%	28%	46%	38%	47%	
Ucayali	41%	24%	13%	41%	59%	41%	68%
Nacional	40%	33%	35%	42%	44%	42%	24%

Fuente: elaboración propia en base a ENAHO – INEI, 2022



En la tabla 5 nos indica que el ingreso promedio de los trabajadores formales e informales varía según el nivel educativo. La mayor brecha se observa en el grupo que tiene secundaria incompleta, donde los formales ganan un 40% más que los informales. Esta diferencia se reduce a un 33% en el grupo con secundaria completa. Los trabajadores con educación superior no universitaria incompleta tienen una brecha del 35%, mientras que los que la tienen completa presentan una brecha del 42%. Los trabajadores con educación superior universitaria incompleta y completa tienen brechas similares, del 44% y 42% respectivamente. Finalmente, el grupo con maestría y/o doctorado tiene la menor brecha, del 24%. Estos resultados indican que la brecha salarial entre formales e informales disminuye a medida que aumenta el nivel educativo.

En el grupo de trabajadores con educación secundaria incompleta, la mayor brecha salarial entre formales e informales se da en Puno, donde los primeros ganan un 77% más que los segundos en promedio. En el grupo de trabajadores con educación secundaria completa, la brecha más alta corresponde a Huánuco, con un 65%, y la más baja a Lima, con un 21%. En el grupo de trabajadores con educación superior no universitaria incompleta, Cajamarca presenta la brecha más elevada, con un 80%, mientras que Madre de Dios tiene la más reducida, con un 10%. En el grupo de trabajadores con educación superior universitaria completa, Huancavelica tiene la brecha más amplia, con un 71%, y Huánuco la más estrecha, con un 15%. En el grupo de trabajadores con educación superior completa, Madre de Dios muestra la brecha más grande, con un 63%, y Huancavelica la más pequeña, con un 2%. Finalmente, en el grupo de trabajadores con maestría y/o doctorado, Ica tiene la brecha más alta, con un 80%, y Loreto la más baja, con un 14%.

Análisis del Número de casos para el nivel educativo

Tabla 6

Número de Casos por Nivel Educativo

Depart.	Secund. Incomp.	Secund. Compl.	Superior no universitaria Incompleta	Superior no universitaria completa	Sup. Univer. Incompl.	Sup. Univers. Comple.	Maestría Doctor.
Amazonas	99%	88%	82%	77%	41%	25%	5%
Ancash	84%	75%	76%	67%	44%	27%	7%
Apurímac	97%	90%	69%	80%	35%	26%	
Arequipa	85%	79%	65%	64%	31%	26%	9%
Ayacucho	93%	87%	78%	80%	37%	28%	8%
Cajamarca	94%	85%	86%	70%	23%	24%	11%
Callao	81%	57%	53%	50%	31%	20%	20%
Cusco	92%	85%	78%	59%	42%	27%	19%
Huancavelica	95%	88%	94%	82%	33%	31%	24%
Huánuco	97%	85%	85%	75%	31%	28%	5%
Ica	61%	62%	57%	52%	34%	29%	6%
Junín	95%	77%	75%	74%	41%	32%	10%
La Libertad	79%	73%	59%	54%	36%	24%	
Lambayeque	81%	69%	74%	67%	42%	31%	9%
Lima	84%	68%	57%	53%	24%	25%	14%
Loreto	90%	75%	63%	64%	26%	19%	8%
Madre de Dios	95%	85%	68%	85%	38%	25%	
Moquegua	79%	60%	51%	63%	30%	30%	12%
Pasco	96%	67%	68%	66%	32%	34%	6%
Piura	83%	76%	67%	66%	37%	24%	5%
Puno	98%	89%	85%	93%	54%	37%	21%
San Martín	93%	79%	83%	69%	34%	29%	5%
Tacna	86%	81%	82%	69%	36%	38%	7%
Tumbes	88%	70%	68%	57%	35%	25%	
Ucayali	88%	76%	66%	70%	27%	33%	15%
Nacional	87%	74%	67%	64%	35%	27%	10%

Fuente: elaboración propia en base a ENAHO – INEI, 2022



Como se puede apreciar en la tabla 6, el nivel educativo influye en la diferencia porcentual entre los trabajadores informales y los formales. La brecha más amplia corresponde al grupo con secundaria incompleta, donde hay un 87% más de trabajadores informales que formales. Esta brecha se reduce gradualmente en los grupos con secundaria completa (74%), educación superior no universitaria incompleta (67%) y completa (64%). Los grupos con educación superior universitaria incompleta y completa presentan brechas menores, del 35% y 27% respectivamente. Finalmente, el grupo con maestría y/o doctorado tiene la brecha más pequeña, del 10%. Estos resultados muestran que la informalidad laboral disminuye a medida que se incrementa el nivel educativo.

El grupo de trabajadores con educación secundaria incompleta tiene la mayor brecha entre formales e informales en Amazonas, donde los primeros son un 99% más que los segundos. La menor brecha se da en Ica, con un 61%. En el grupo con secundaria completa, Apurímac tiene la brecha más alta (90%) y el Callao la más baja (57%). En el grupo con educación superior no universitaria incompleta, la brecha más elevada es en Huancavelica (94%) y la más reducida en Moquegua (51%). En el grupo con educación superior no universitaria completa, Puno tiene la brecha más amplia (54%) y Cajamarca la más estrecha (23%). En el grupo con educación superior incompleta, Puno muestra la brecha más grande (93%) y el Callao la más pequeña (50%). En el grupo con educación superior completa, Tacna tiene la brecha más grande (38%) y Loreto la más pequeña (19%). En el grupo con maestría y/o doctorado, Ica tiene la brecha más alta (24%) y Amazonas, Huánuco y Piura tienen la brecha más baja (5%).

4.1.2. Resultados del análisis econométrico

Para el Objetivo 1

Para cumplir con el objetivo 1, se realizó un análisis econométrico de los datos. Se estimaron dos modelos Logit y dos Probit para examinar la relación entre la variable dependiente y las variables independientes.

Tabla 7

Criterios estadísticos y valores para seleccionar el modelo más adecuado

Variab les	Logit	Probit
Observaciones	54582	54582
Chi_2	11350.058	14383.189
Pseudo - R2	0.41986517	0.42025964
ll	-16708.665	-16697.304

Fuente: elaboración propia en base a resultado de Stata 17 con datos ENAHO – INEI, 2022

De acuerdo a Gujarati y Porter (2010), se selecciona el modelo más adecuado para los datos según los siguientes criterios, mostrados en la tabla 7:

Chi cuadrado (Chi_2): Esta es una medida de la bondad de ajuste del modelo. Un valor más alto indica que el modelo se ajusta mejor a los datos. En los resultados, el valor de Chi cuadrado es mayor en el modelo probit que en el modelo logit, lo que sugiere que el modelo probit se ajusta mejor a tus datos.

Pseudo R-cuadrado (Pseudo - R2): Es un indicador de la bondad de ajuste para modelos logit y probit. No se puede interpretar como el R-cuadrado de la regresión lineal, que mide la proporción de la variabilidad explicada por el modelo. Pero sirve para comparar diferentes modelos con los mismos datos. En los resultados, los Pseudo R-cuadrados de ambos modelos son parecidos, lo que

indica que explican una cantidad similar de la variabilidad en los datos, con una leve ventaja para el modelo probit.

Log-verosimilitud (ll): Este es el logaritmo de la función de verosimilitud evaluada en los estimadores de máxima verosimilitud. En términos simples, es una medida de cuán probable es el modelo dado los datos. Un valor más alto (más cercano a cero) indica un mejor ajuste del modelo. En los resultados, el valor de la log-verosimilitud es ligeramente mayor (más cercano a cero) en el modelo probit que en el modelo logit, lo que sugiere que el modelo probit se ajusta ligeramente mejor a los datos.

Tabla 8

Resultados del Modelo Logit y Probit estimado

Variab les	Logit	Probit
edu	-.25934751***	-.137398***
exp	-.09123891***	-.04796368***
exp2	.001059***	.00052277***
sexo	.25307684***	.11520748***
nemp	-.00005526***	-.00002819***
temp	-.81838663***	-.47529762***
_cons	6.9054808***	3.8409763***

Legend: *p<.1; **p<.05; *** p<.01

Fuente: elaboración propia en base a resultado de Stata 17 con datos ENAHO – INEI, 2022

Los coeficientes de los modelos logit y probit indican cómo varía el logaritmo de las probabilidades o la función de distribución acumulativa de la distribución normal estándar, respectivamente, cuando una variable aumenta en una unidad y el resto se mantiene constante. Esto se explica en Gujarati y Porter (2010) y Wooldridge (2020).



Ahora, vamos a interpretar los coeficientes de cada variable:

EDU (Años de educación alcanzados): En ambos modelos, los coeficientes son negativos y significativos, lo que indica que a medida que los años de educación aumentan, la probabilidad de participar en el mercado laboral de manera informal disminuye.

EXP (Experiencia potencial): Los coeficientes son negativos y significativos en ambos modelos, lo que sugiere que a medida que la experiencia potencial aumenta, la probabilidad de participar en el mercado laboral de manera informal disminuye.

EXP² (Experiencia potencial al cuadrado): Los coeficientes son positivos y significativos en ambos modelos, lo que indica que hay un efecto cuadrático de la experiencia en la probabilidad de participar en el mercado laboral de manera informal. Esto podría sugerir que después de cierto punto, el aumento de la experiencia podría aumentar la probabilidad de participar en el mercado laboral de manera informal.

SEXO (Sexo): Los coeficientes son positivos y significativos en ambos modelos, lo que indica que ser de sexo masculino aumenta la probabilidad de participar en el mercado laboral de manera informal.

NEMP (Número de nuevas empresas): Los coeficientes son negativos y significativos en ambos modelos, lo que sugiere que a medida que el número de nuevas empresas aumenta, la probabilidad de participar en el mercado laboral de manera informal disminuye.

TEMP (Número de trabajadores en la empresa): Los coeficientes son negativos y significativos en ambos modelos, lo que indica que a medida que el número de trabajadores en la empresa aumenta, la probabilidad de participar en el mercado laboral de manera informal disminuye.

Tabla 9

Efectos marginales del modelo estimado

Variab les	Logit_mfx	Probit_mfx
edu	-0.0285826	-0.0303568
exp	-0.0100554	-0.0105971
exp2	0.0001167	0.0001155
sexo	0.0278915	0.025454
nwemp	-6.09E-06	-6.23E-06
tmemp	-0.0901942	-0.1050125

Fuente: elaboración propia en base a resultado de Stata 17 con datos ENAHO – INEI, 2022

La tabla 9 nos muestra los resultados de los efectos marginales, estos representan el cambio en la probabilidad de que la variable dependiente sea igual a 1 (en este caso, que la participación en el mercado laboral sea informal) para un cambio unitario en la variable correspondiente, manteniendo constantes las demás variables.

Ahora, vamos a interpretar los efectos marginales de cada variable:

EDU (Años de educación alcanzados): En ambos modelos, los efectos marginales son negativos, lo que indica que a medida que los años de educación aumentan en una unidad, la probabilidad de participar en el mercado laboral de manera informal disminuye en un 2.86% en el modelo logit y en un 3.04% en el modelo probit.



EXP (Experiencia potencial): Los efectos marginales son negativos en ambos modelos, lo que sugiere que a medida que la experiencia potencial aumenta en una unidad, la probabilidad de participar en el mercado laboral de manera informal disminuye en un 1.01% en el modelo logit y en un 1.06% en el modelo probit.

EXP² (Experiencia potencial al cuadrado): Los efectos marginales son positivos en ambos modelos, lo que indica que hay un efecto cuadrático de la experiencia en la probabilidad de participar en el mercado laboral de manera informal. Esto podría sugerir que después de cierto punto, el aumento de la experiencia podría aumentar la probabilidad de participar en el mercado laboral de manera informal en un 0.0117% en el modelo logit y en un 0.01155% en el modelo probit por cada unidad adicional de experiencia al cuadrado.

SEXO (Sexo): Los efectos marginales son positivos en ambos modelos, lo que indica que ser de sexo masculino aumenta la probabilidad de participar en el mercado laboral de manera informal en un 2.79% en el modelo logit y en un 2.55% en el modelo probit.

NEMP (Número de nuevas empresas): Los efectos marginales son negativos en ambos modelos, lo que sugiere que a medida que el número de nuevas empresas aumenta en una unidad, la probabilidad de participar en el mercado laboral de manera informal disminuye en un 0.000609% en el modelo logit y en un 0.000623% en el modelo probit.

TEMP (Número de trabajadores en la empresa): Los efectos marginales son negativos en ambos modelos, lo que indica que a medida que el nivel de trabajadores en la empresa aumenta en una unidad, la probabilidad de participar

en el mercado laboral de manera informal disminuye en un 9.02% en el modelo logit y en un 10.50% en el modelo probit.

Para el Objetivo 2

Para cumplir con el objetivo 2, se realizó un análisis econométrico de los datos. Se estimaron dos modelos para la obtención de brechas salariales entre empleo formal e informal.

Análisis de brecha de ingresos metodología Oaxaca – Blinder:

Tabla 10

Resultados obtenidos de la metodología Oaxaca - Blinder

ingm	Coefficient	Std. err.	z	P> z 	[95% conf. interval]	
overall						
group_1	7.594967	0.0057118	1329.7	0	7.583772	7.606162
group_2	6.624896	0.0068179	971.7	0	6.611533	6.638259
difference	0.970071	0.0088943	109.07	0	0.9526386	0.9875035
endowments	0.5979252	0.0142582	41.94	0	0.5699797	0.6258707
coefficients	0.4855345	0.0120359	40.34	0	0.4619446	0.5091244
interaction	-0.1133887	0.0163624	-6.93	0	-0.1454584	-0.0813189

Fuente: elaboración propia en base a resultado de Stata 17 con datos ENAHO – INEI, 2022

Siguiendo a Oaxaca (1973) y Blinder (1973), procederemos a interpretar los resultados obtenidos:

Diferencia total (overall difference): La diferencia total en los ingresos laborales mensuales (en logaritmos) entre los dos grupos es de 0.970071. Esto significa que, en promedio, el grupo 1 (empleo formal) tiene ingresos laborales mensuales (en logaritmos) 0.970071 unidades más altos que el grupo 2 (empleo informal).



Parte explicada (endowments): De la diferencia total, 0.5979252 se puede explicar por las diferencias en las características observables, entre los dos grupos. Esto significa que si el grupo 2 (empleo informal) tuviera las mismas características observables que el grupo 1 (empleo formal), la diferencia en los ingresos laborales mensuales (en logaritmos) se reduciría en 0.5979252.

Parte no explicada (coefficients): De la diferencia total, 0.4855345 no se puede explicar por las diferencias en las características observables y se atribuye a las diferencias en los coeficientes. Esto podría deberse a la discriminación, a las diferencias en las características no observables, o a la especificación del modelo.

Interacción (interaction): El término de interacción (-0.1133887) captura el efecto de la interacción entre las diferencias en las características observables y las diferencias en los coeficientes. En este caso, el término de interacción es negativo, lo que significa que reduce la diferencia total en los ingresos laborales mensuales (en logaritmos) entre los dos grupos.

Análisis de brecha de ingresos metodología Ñopo:

Tabla 11

Resultados obtenidos de la Metodología Ñopo

Variables	Coeficiente
D	-0.12772551
D0	-0.04699268
DM	-0.02657566
DF	-0.01856749
DX	-0.03558968
percM	0.31949434
percF	0.31405137

Fuente: elaboración propia en base a resultado de Stata 17 con datos ENAHO – INEI, 2022



Para analizar la brecha de ingresos entre dos grupos, se aplica la metodología de Ñopo (2004), que permite separar la parte atribuible a las diferencias en las características observables (endowments) de la parte no atribuible (discriminación). Los resultados se interpretan de la siguiente manera:

D: Este es el coeficiente de la brecha total en los ingresos laborales mensuales (en logaritmos) entre los dos grupos. Un valor de -0.12772551 indica que el grupo de referencia (empleo informal) tiene ingresos laborales mensuales (en logaritmos) 0.12772551 unidades más bajos que el otro grupo (empleo informal), o su equivalente a que el empleo formal tiene 0.8722744 unidades más altos en ingreso que empleo informal.

D0: Este es el coeficiente de la brecha en los ingresos que no se puede explicar por las diferencias en las características observables entre los dos grupos. Un valor de -0.04699268 indica que incluso si el grupo de referencia (empleo informal) tuviera las mismas características observables que el otro grupo (empleo formal), todavía habría una brecha de ingresos de 0.04699268 .

DM: Este es el coeficiente de la brecha en los ingresos que se puede explicar por las diferencias en las características observables entre dos grupos de individuos con $by\text{-variable}=1$, que están dentro y fuera del soporte común de las variables definidas por $match\text{-variable}(s)$. Un valor de -0.02657566 indica que si los individuos con $by\text{-variable}=1$ del grupo de referencia (empleo formal) tuvieran las mismas características observables que los individuos con $by\text{-variable}=1$ del otro grupo (empleo informal), la brecha de ingresos se aumentaría en 0.02657566 .

DF: Este es el coeficiente de la brecha en los ingresos que se puede explicar por las diferencias en las características observables entre dos grupos de

individuos con $by\text{-variable}=0$, que están dentro y fuera del soporte común de las variables definidas por $match\text{-variable}(s)$. Un valor de -0.01856749 indica que si los individuos con $by\text{-variable}=0$ del grupo de referencia (empleo formal) tuvieran las mismas características observables que los individuos con $by\text{-variable}=0$ del otro grupo (empleo informal), la brecha de ingresos se reduciría en 0.01856749 .

DX: Este es el coeficiente de la brecha en los ingresos que se puede explicar por las diferencias en la distribución de las características de los dos grupos discriminados por $by\text{-variable}$ en el soporte común de las variables definidas por $match\text{-variable}(s)$. Un valor de -0.03558968 indica que si el grupo de referencia (empleo formal) tuviera la misma distribución de características que el otro grupo (empleo informal), la brecha de ingresos se reduciría en 0.03558968 .

4.2. DISCUSIÓN

El objetivo 1 muestra que la educación (EDU), en años alcanzados, tiene un efecto significativo en la probabilidad de ser informal en el mercado laboral. Los dos modelos tienen efectos marginales negativos, lo que significa que un año más de educación reduce la probabilidad de informalidad en un 2.86% (modelo logit) y 3.04% (modelo probit). La experiencia potencial (EXP) también influye negativamente en esta probabilidad, indicando que un aumento de una unidad en EXP disminuye la probabilidad de informalidad en un 1.01% (modelo logit) y 1.06% (modelo probit). Sin embargo, la variable EXP^2 (Experiencia potencial al cuadrado) muestra un efecto cuadrático, lo que implica que después de cierto nivel, la experiencia podría aumentar levemente la probabilidad de informalidad en un 0.0117% (modelo logit) y 0.01155% (modelo probit) por cada unidad adicional de EXP^2 . En relación al género (SEXO), ser hombre se asocia con una



mayor probabilidad de ser informal, con efectos marginales positivos del 2.79% (modelo logit) y 2.55% (modelo probit). Por último, las variables de creación de nuevas empresas (NEMP) y número de trabajadores por empresa (TEMP) tienen efectos marginales negativos, lo que significa que un incremento de una unidad en NEMP y TEMP se relaciona con una menor probabilidad de informalidad laboral, con efectos marginales de 0.000609% y 0.000623% en el modelo logit y probit, respectivamente, para NEMP, y de 9.02% y 10.50% para TEMP.

En diversas investigaciones, abordadas en los antecedentes, se han identificado algunos factores que influyen en la probabilidad de estar en el sector informal laboral, tales como el nivel educativo, el género, la edad y el tamaño de la empresa. Por ejemplo, Carvajal et al. (2017), encontraron que un mayor nivel educativo, el género masculino y una mayor edad reducen la probabilidad de informalidad. Ramírez et al. (2015), reportaron que las mujeres tienen un 11.73% más de probabilidad de estar en la informalidad que los hombres. Solis (2021), estimó que, al incrementar el nivel educativo, la probabilidad de informalidad se reduce en un 20.09%, mientras que el sexo femenino y el aumento de la edad se asocian con aumentos en la probabilidad de informalidad. Ramos (2020), mostró que un aumento en educación reduce la probabilidad de informalidad en un 3%, así como la edad y el tamaño de la empresa disminuyen la probabilidad de participación laboral en el sector informal en un 2.8% y 4%, respectivamente. Machaca (2019), señaló que la informalidad está afectada por factores como el nivel educativo (-6%), la experiencia laboral (-0.5%), el género y el tamaño de la empresa (78%) y por último tenemos a Pedraza (2016), que nos dice que una menor probabilidad de empleo informal se relaciona con un mayor nivel educativo y mayor edad.



El Objetivo 2, Según los datos de ENAHO 2022 y el método Oaxaca – Blinder, hay una brecha salarial importante entre formales e informales. Siendo este la diferencia total en los ingresos mensuales en logaritmos de 0.97. Lo que equivale que los formales ganan 0.97 unidades logarítmicas más, que los informales, ajustando por varios factores. La brecha se debe en parte a diferencias observables como educación, experiencia, número de nuevas empresas y número de trabajadores en la empresa (0.5979252). Si los informales tuvieran estas características, la brecha se reduciría en ese porcentaje. La otra parte se debe a diferencias no observables o en coeficientes, que pueden reflejar discriminación, características ocultas o la especificación del modelo (0.4855345). El término de interacción (-0.1133887) tiene un efecto negativo, disminuyendo la diferencia total en ingresos mensuales entre ambos grupos.

Por otro lado, La metodología de Ñopo (2004) ofrece la ventaja de construir una distribución salarial flexible, sin imponer una forma funcional que limite el comportamiento de las variables aleatorias. Además, permite desglosar la brecha salarial en sus componentes, en particular entre trabajadores formales e informales con características similares, lo que aporta más información que solo examinar el promedio. Según esta metodología, se puede contrastar un conjunto de trabajadores formales con otro de informales, basándose en características comunes, mediante una técnica de emparejamiento. Este emparejamiento se denomina soporte común¹³ y se divide en cuatro subgrupos: i) formales dentro del soporte común; ii) formales fuera del soporte común; iii) informales dentro del soporte común; e iv) informales fuera del soporte común. La estimación para el 2022, muestra que el ingreso laboral de los trabajadores formales fue 0.872 unidades logarítmicas, más que al de los trabajadores informales, es decir, que el



beneficio de la formalidad sería bastante alto en comparación con los informales, considerando las características como educación, experiencia, número de nuevas empresas y número de trabajadores en la empresa.

Asimismo, el ingreso de los trabajadores formales e informales en América Latina presenta una brecha significativa, según diversos estudios que aplican la metodología de Oaxaca – Blinder. Por ejemplo, Maurizio (2014) reporta diferencias salariales promedio de alrededor del 34% en Argentina y Ecuador, el 28% en Uruguay y cerca del 20% en los demás países de América Latina, incluyendo el 24.6% en el Perú. Carrillo (2004) también estima una brecha de 18% para el Ecuador, a favor de los formales. Esparta y Rivera (2020) hallan una brecha de 87.7% para el Perú con la misma metodología, y de 100% con la metodología de Ñopo. Finalmente, MACHACA (2019) calcula una brecha de 0.867 unidades para el Perú con Oaxaca y Blinder, favoreciendo a los formales.



V. CONCLUSIONES

- El Perú enfrenta el desafío de la informalidad laboral, que afecta al 76.1% de los trabajadores peruanos y al 61.2% de los trabajadores mundiales. Esta situación se agrava en regiones como Puno, donde la informalidad supera en un 70% a la formalidad, o en Huancavelica, Amazonas, Apurímac y Madre de Dios, donde la brecha es de 54%. Asimismo, se ha comprobado que factores como la educación y la creación de nuevas empresas influyen en el empleo informal y en la diferencia de ingresos entre los ocupados informales y formales, según un modelo que considera otras variables como la experiencia, el sexo y el número de trabajadores en la empresa, para el periodo y la población analizados.
- El nivel educativo y la creación de empresas tienen una influencia importante en la posibilidad de ser trabajador informal. Según los dos modelos, cada año adicional de educación disminuye la posibilidad de informalidad en un 2.86% (modelo logit) y 3.04% (modelo probit). Asimismo, para la creación de nuevas empresas (NEMP), y de 9.02% y 10.50% para el número de empleados por empresa (TEMP).
- De acuerdo con el método Oaxaca – Blinder, los trabajadores formales e informales tienen una diferencia significativa en sus ingresos por hora. Los formales perciben 0.97 unidades logarítmicas más que los informales, lo que implica un mayor ingreso mensual. Esta diferencia se debe a factores observables como el nivel educativo, la experiencia laboral, el número de empresas nuevas y el tamaño de la empresa (0.5979252), que benefician a los formales. Si los informales contaran con estas características, la brecha se disminuiría en esa medida. Sin embargo, también existen factores no observables o de coeficientes (0.4855345), que pueden estar relacionados con discriminación, variables omitidas o la especificación del modelo. Además, esta brecha se calculó con la metodología de Ñopo, que también muestra una brecha



salarial de 0.872 unidades logarítmicas a favor de los formales en el 2022,
considerando las mismas características.



VI RECOMENDACIONES

- Mejorar el nivel educativo de las personas, que es el factor más importante para disminuir el empleo informal.
- Capacitar o fomentar la experiencia laboral de las personas, que también influye positivamente en la reducción de la informalidad.
- Impulsar la formalización de las microempresas y generar políticas que favorezcan su crecimiento, para que puedan contratar a más trabajadores formales. Esto es clave, ya que el tamaño de la empresa es el principal determinante del empleo informal. Para ello, se podrían reducir los impuestos escalonados y simplificar los trámites burocráticos para la formalización.



VII. REFERENCIAS BIBLIOGRÁFICAS

- Instituto Nacional de Estadística e Informática. (2017). Características de la Población Ocupada en Empleo Formal e Informal. *Characteristics of the Population Employed in Formal and Informal Employment*.
https://www.inei.gob.pe/media/MenuRecursivo/publicaciones_digitales/Est/Lib1537/cap07.pdf
- Instituto Nacional de Estadística e Informática. (2017). Definiciones y conceptos. *Departamento Administrativo Nacional de Estadística*, 147–148.
https://www.inei.gob.pe/media/MenuRecursivo/publicaciones_digitales/Est/Lib1676/06.pdf
- Instituto Nacional de Estadística e Informática. (2022). Informe Técnico Empleo Nacional Ene-Feb-Mar-2022. *Informe Técnico, 0*, 1–40. <https://bit.ly/3PofxYU>
- Ministerio de Economía y Finanzas. (2023). *Contenido Panorama internacional*.
https://cdn.www.gob.pe/uploads/document/file/4486854/PPT_IAPM_27042023.pdf
- Adanaqué Mamani, A. (2022). *Educación y desigualdad salarial en el Perú* [Universidad Nacional Mayor de San Marcos].
https://cybertesis.unmsm.edu.pe/bitstream/handle/20.500.12672/19665/Adanaque_ma.pdf?sequence=1&isAllowed=y
- Alvarado, A. (2018). Empleo Informal Y Mercados Laborales Multisegmentados en el Perú. [Pontificia Universidad Católica del Perú]. In *Pontificia Universidad Católica Del Perú*.
<https://tesis.pucp.edu.pe/repositorio/handle/20.500.12404/12366>
- Arpi Mayta, Quilca, R., & Arpi Quilca, L. (2010). DESIGUALDAD DEL INGRESO LABORAL Y NIVEL EDUCATIVO ENTRE GRUPOS ÉTNICOS EN EL PERÚ. *Scielo*, 44(9), 56–67.
http://www.scielo.org.pe/scielo.php?script=sci_arttext&pid=S2219-71682018000100006
- Barco, D., & Vargas, P. (2010). Brechas Salariales entre Formales e Informales. *XXVII*



Encuentro de Economistas, Mayo, 14.

<https://www.bcrp.gob.pe/docs/Publicaciones/Documentos-de-Trabajo/2010/Documento-de-Trabajo-03-2010.pdf>

Becker, G. S. (1975). Human Capital: A theoretical and empirical analysis, with special reference to Education. *Archives of Neurology*, 43(1), 58–61.

<https://www.nber.org/system/files/chapters/c3730/c3730.pdf>

BILLOROU GARZON, J. C. (2021). *Diferencias en el ingreso entre la mujer Urbana y Rural en Colombia* [UNIVERSIDAD COLEGIO MAYOR DE CUNDINAMARCA].

https://repositorio.unicolmayor.edu.co/bitstream/handle/unicolmayor/5501/AS002L_1.PDF?sequence=1&isAllowed=y

Blanchard, O., Amighini, A., & Giavazzi, F. (2012). *Macroeconomía*. PEARSON EDUCACIÓN, S.A. 2da edición.

Blinder, A. S. (1973). Wage Discrimination: Reduced Form and Structural ESTimates. In *The journal of human resources* (Vol. 8, Issue 4, pp. 436–455).

<https://home.cerge-ei.cz/munich/ecm04/in/blinder.pdf>

Calatayud Mendoza, A. P. (2019). Brecha Salarial Por Género De Los Trabajadores Dependientes En El Perú 2012-2016. *Semestre Económico*, 7(2), 40–54.

<https://doi.org/10.26867/seconomico.v7i2.226>

Carrillo, P. (2004). Las diferencias salariales entre el sector público y privado en el Ecuador. *Cuestiones Económicas*, 20(2), 165–175.

<https://repositorio.bce.ec/bitstream/32000/232/1/XX-II-05CARRILLO.pdf>

Carvajal Calderón, A. N., Cárdenas Hernández, L. Y., & Estrada Cañas, I. (2017). Determinantes socioeconómicos de la informalidad laboral y el subempleo en el área metropolitana de Bucaramanga, Colombia. *Equidad y Desarrollo*, 1(29), 53–82. <https://doi.org/10.19052/ed.4171>

Casarreal, J., & Cruz, M. (2020). Empleo informal: una explicación desde la demanda. *Contaduría y Administración*, 66(1), 239.

<https://doi.org/10.22201/fca.24488410e.2021.2595>



- Esparta, D., & Rivera, G. (2020). Brechas salariales entre el empleo formal e informal. *Boletín Informativo Laboral*, 104, 1–18.
<https://cdn.www.gob.pe/uploads/document/file/1297314/Art 2 - Brechas salarial.pdf>
- Galiani, S., & Weinschelbaum, F. (2007). *Modeling Informality Formally: Households and Firms Documento* (No. 47; Issue 47).
https://www.econstor.eu/bitstream/10419/189602/1/doc_cedlas47.pdf
- Garavito, C. (2011). Desigualdad en los ingresos : género y lengua materna. In *Desigualdad distributiva en el Perú: dimensiones* (pp. 235–265). Pontificia Universidad Católica del Perú. <https://doi.org/10.18800/9789972429743.006>
- Gujarati, D. N., & Porter, D. C. (2010). *ECONOMETRÍA* (5ta ed.).
- Hernández, K. C., Contreras, M., & Quejada, R. (2014). Determinantes de la Informalidad Laboral: Un Análisis para Colombia. *Investigación y Desarrollo*, 22(510), 2–22. <https://search-ebshost-com.bibliotecavirtualunap.remotexs.co/login.aspx?direct=true&db=fua&AN=95806027&lang=es&site=ehost-live>
- Hernández Sampieri, R., Fernández Collado, C., & Baptista Lucio, M. del P. (2023). METODOLOGÍA DE LA INVESTIGACIÓN CIENTÍFICA Y EDUCATIVA. In *Metodología de la investigación científica y educativa* (Quinta, pp. 9–34). Atena Editora. <https://doi.org/10.22533/at.ed.6962318092>
- Instituto Nacional de Estadística e Informática [INEI]. (2020a). *Producción y Empleo Informal en el Perú*.
<https://cdn.www.gob.pe/uploads/document/file/3441492/Producción y Empleo Informal en el Perú%2C Cuenta Satélite de la Economía Informal 2007-2019.pdf?v=1658438890>
- Instituto Nacional de Estadística e Informática [INEI]. (2020b). *Reporte de la Encuesta nacional de hogares (ENAHO)*.
- Krugman, P., & Wells, R. (2009). *MICROECONOMICS*. 5ta edición.
- MACHACA, L. E. (2019). *DETERMINANTES DE LA INFORMALIDAD LABORAL Y*



*BRECHAS DE INGRESOS ENTRE EMPLEOS FORMALES E INFORMALES
EN EL PERÚ, PERIODO 2017* [Universidad Nacional del Altiplano].

http://repositorio.unap.edu.pe/bitstream/handle/20.500.14082/12453/Machaca_Flores_Lizbeth_Estefany.pdf?sequence=1&isAllowed=y

Maurizio, R. (2014). Informalidad laboral y brechas salariales en América Latina [Universidad Nacional de General Sarmiento y CONICET]. In *Biblioteca Jurídica Virtual del Instituto de Investigaciones Jurídicas de la UNAM*.
www.juridicas.unam.mx<http://biblio.juridicas.unam.mx>

Mendoza Bellido, W. (2014). *COMO INVESTIGAN LOS ECONOMISTAS*. Pontificia Universidad Católica del Perú.

Mincer, J. (1974). Education, Experience and the Distribution of Earnings and Employment. In *Southern Economic Journal* (Vol. 43, Issue 3, p. 1407). NBER.
<https://www.nber.org/system/files/chapters/c3693/c3693.pdf><https://www.nber.org/system/files/chapters/c3693/c3693.pdf>

Nicholson, W. (2008). *Teoría Microeconómica*. 2da edición
<https://www.nber.org/system/files/chapters/c3730/c3730.pdf>

Ñopo, H. (2004). Matching as a Tool to Decompose Wage Gaps. *Review of Economics and Statistics*, 90(2), 290–299. <https://doi.org/10.1162/rest.90.2.290>

Oaxaca, R. (1973). Male - Female Wage Differentials in Urban Labor Markets. In *International Economic Review*.
https://inequality.stanford.edu/sites/default/files/media/_media/pdf/Classic_Media/Oaxaca_1973_Discrimination_and_Prejudice.pdf

Organización Internacional del Trabajo. (2020). Promover el empleo y el trabajo decente en un panorama cambiante. In *Informe III (Parte B)*.
https://www.ilo.org/ilc/ILCSessions/109/reports/reports-to-the-conference/WCMS_738283/lang--es/index.htm

Organización Internacional del Trabajo. (2017). El trabajo decente y la Agenda 2030 de Desarrollo Sostenible. *Gestión Prácticas de Riesgo Laborales: Integración y Desarrollo de La Gestión de La Prevención*, 10–11.
<https://dialnet.unirioja.es/servlet/articulo?codigo=7877986>



- Ovando-Aldana, W., Rivera-Rojo, C. R., & Salgado-Vega, M. C. (2021). Características del empleo informal en México, 2005 y 2020. *Papeles de Población*, 27(108), 147–184. <https://doi.org/10.22185/24487147.2021.108.15>
- Paz, J. A. (2021). La brecha de remuneraciones entre segmentos del mercado de trabajo en la Argentina. *Revista Desarrollo y Sociedad*, 2021(88), 239–271. <https://doi.org/10.13043/DYS.88.7>
- Pedraza Almiron, T. (2016). *Factores que inciden en el empleo informal en el departamento del cusco, en el año 2014* [Universidad Andina del Cusco]. http://repositorio.uandina.edu.pe/bitstream/UAC/803/1/Tatiana_Tesis_bachiller_2016.pdf
- Ramírez Gallego, J. B., Avila Carreño, C. A., & Arias Manrique, I. J. (2015). Factores que inciden en la probabilidad de permanecer en la informalidad en Colombia (2008-2012): un análisis de las medidas de política pública. *Revista Facultad de Ciencias Económicas*, 23(2), 9–20. <https://doi.org/10.18359/rfce.1604>
- RAMOS, J. M. (2020). *DETERMINANTES DEL EMPLEO INFORMAL EN EL DEPARTAMENTO DE PUNO PARA EL PERIODO 2008 - 2017* [Universidad Nacional del Altiplano]. http://repositorio.unap.edu.pe/bitstream/handle/20.500.14082/15136/Ramos_Ito_Jhon_Mauricio.pdf?sequence=1&isAllowed=y
- Rentería, J. M. (2015). BRECHAS DE INGRESOS LABORALES EN EL PERÚ URBANO: UNA EXPLORACIÓN DE LA ECONOMÍA INFORMAL. In *Departamento De Economía-Pucp* (No. 408; Issue ¿POR QUÉ SE DAN LAS BRECHAS DE INGRESOS LABORALES EN EL PERÚ URBANO?). <https://files.pucp.education/departamento/economia/DDD408.pdf>
- Sánchez - Villagómez, M., & Chafloque - Céspedes, R. (2019). *LA INFORMALIDAD LABORAL EN EL PERÚ* (F. E. USMP (ed.)). <https://repositorio.usmp.edu.pe/handle/20.500.12727/5680>
- Solis, J. (2021). Factores que influyen en la informalidad laboral desde la perspectiva de la oferta en la provincia de Lima, Perú, en el año 2018 [Universidad ESAN]. In *Pesquisa Veterinaria Brasileira* (Vol. 26, Issue 2).



[https://repositorio.esan.edu.pe/bitstream/handle/20.500.12640/2749/2021_ECY
NI_21-2_01_T.pdf?sequence=1&isAllowed=y](https://repositorio.esan.edu.pe/bitstream/handle/20.500.12640/2749/2021_ECY_NI_21-2_01_T.pdf?sequence=1&isAllowed=y)

Wooldridge, J. M. (2020). *Introducción a la econometría: Un enfoque moderno* (4a.,
Issue July). Cengage Learning Edi to res, S.A. <http://latinoamerica.cengage.com>



ANEXOS

ANEXO A. Fuente de Obtención de datos

CONSULTA POR ENCUESTA

Sírvase seleccionar Encuesta, Año y Período y a continuación se mostrarán todos los Módulos de la Encuesta Seleccionada. Luego proceda a descargar el módulo de su interés.

ENCUESTA ENAHO Metodología ACTUALIZADA

Condiciones de Vida y Pobreza - ENAHO

AÑO 2022 Período: Anual -(Ene-Dic)

Nro	Año	Período	Código Encuesta	Encuesta	Código Módulo	Módulo	Ficha	Descarga
1	2022	55	784	Condiciones de Vida y Pobreza - ENAHO	1	Características de la Vivienda y del Hogar		
2	2022	55	784	Condiciones de Vida y Pobreza - ENAHO	2	Características de los Miembros del Hogar		
3	2022	55	784	Condiciones de Vida y Pobreza - ENAHO	3	Educación		
4	2022	55	784	Condiciones de Vida y Pobreza - ENAHO	4	Salud		
5	2022	55	784	Condiciones de Vida y Pobreza - ENAHO	5	Empleo e Ingresos		
6	2022	55	784	Condiciones de Vida y Pobreza - ENAHO	7	Gastos en Alimentos y Bebidas (Módulo 601)		
7	2022	55	784	Condiciones de Vida y Pobreza - ENAHO	8	Instituciones Benéficas		
8	2022	55	784	Condiciones de Vida y Pobreza - ENAHO	9	Mantenimiento de la Vivienda		
9	2022	55	784	Condiciones de Vida y Pobreza - ENAHO	10	Transportes y Comunicaciones		
10	2022	55	784	Condiciones de Vida y Pobreza - ENAHO	11	Servicios a la Vivienda		
11	2022	55	784	Condiciones de Vida y Pobreza - ENAHO	12	Esparcimiento , Diversion y Servicios de Cultura		
12	2022	55	784	Condiciones de Vida y Pobreza - ENAHO	13	Vestido y Calzado		
13	2022	55	784	Condiciones de Vida y Pobreza - ENAHO	15	Gastos de Transferencias		
14	2022	55	784	Condiciones de Vida y Pobreza - ENAHO	16	Muebles y Enseres		
15	2022	55	784	Condiciones de Vida y Pobreza - ENAHO	17	Otros Bienes y Servicios		
16	2022	55	784	Condiciones de Vida y Pobreza - ENAHO	18	Equipamiento del Hogar		
17	2022	55	784	Condiciones de Vida y Pobreza - ENAHO	22	Producción Agrícola		
18	2022	55	784	Condiciones de Vida y Pobreza - ENAHO	23	Subproductos Agrícolas		
19	2022	55	784	Condiciones de Vida y Pobreza - ENAHO	24	Producción Forestal		
20	2022	55	784	Condiciones de Vida y Pobreza - ENAHO	25	Gastos en Actividades Agrícolas y/o Forestales		
21	2022	55	784	Condiciones de Vida y Pobreza - ENAHO	26	Producción Pecuaria		
22	2022	55	784	Condiciones de Vida y Pobreza - ENAHO	27	Subproductos Pecuarios		
23	2022	55	784	Condiciones de Vida y Pobreza - ENAHO	28	Gastos en Actividades Pecuarias		
24	2022	55	784	Condiciones de Vida y Pobreza - ENAHO	34	Sumarias (Variables Calculadas)		
25	2022	55	784	Condiciones de Vida y Pobreza - ENAHO	37	Programas Sociales (Miembros del Hogar)		

Fuente: INEI -2022

ANEXO B. Resultado de estimación logit

```
. logit ocupinf edu exp exp2 sexo nwemp tmemp, robust
```

```
Iteration 0: log pseudolikelihood = -28801.349
Iteration 1: log pseudolikelihood = -17839.055
Iteration 2: log pseudolikelihood = -16771.716
Iteration 3: log pseudolikelihood = -16708.709
Iteration 4: log pseudolikelihood = -16708.665
Iteration 5: log pseudolikelihood = -16708.665
```

Logistic regression

Number of obs = 54,582
Wald chi2(6) = 11350.06
Prob > chi2 = 0.0000
Pseudo R2 = 0.4199

Log pseudolikelihood = -16708.665

ocupinf	Robust		z	P> z	[95% conf. interval]	
	Coefficient	std. err.				
edu	-.2593475	.0043539	-59.57	0.000	-.267881	-.250814
exp	-.0912389	.0040962	-22.27	0.000	-.0992673	-.0832105
exp2	.001059	.0000795	13.32	0.000	.0009031	.0012149
sexo	.2530768	.0292462	8.65	0.000	.1957554	.3103982
nwemp	-.0000553	6.77e-06	-8.17	0.000	-.0000685	-.000042
tmemp	-.8183866	.0099286	-82.43	0.000	-.8378464	-.7989269
_cons	6.905481	.0826902	83.51	0.000	6.743411	7.067551

Fuente: Elaboración propia en base a datos de ENAHO-2022

ANEXO C. Resultados de efectos marginales del modelo logit

```
. mfx
```

```
Marginal effects after logit
y = Pr(ocupinf) (predict)
= .8738853
```

variable	dy/dx	Std. err.	z	P> z	[95% C.I.]	X
edu	-.0285826	.00044	-64.87	0.000	-.029446 - .027719	9.71146
exp	-.0100554	.00044	-22.86	0.000	-.010918 - .009193	25.9618
exp2	.0001167	.00001	13.64	0.000	.0001 .000133	918.929
sexo	.0278915	.00318	8.78	0.000	.021664 .034119	1.47197
nwemp	-6.09e-06	.00000	-8.20	0.000	-7.5e-06 -4.6e-06	1305.63
tmemp	-.0901942	.0016	-56.41	0.000	-.093328 - .08706	1.65677

Fuente: Elaboración propia en base a datos de ENAHO-2022

ANEXO D. Resultados estimación modelo probit

```
. probit ocupinf edu exp exp2 sexo nwemp tmemp, robust
```

```
Iteration 0: log pseudolikelihood = -28801.349
Iteration 1: log pseudolikelihood = -16920.783
Iteration 2: log pseudolikelihood = -16698.735
Iteration 3: log pseudolikelihood = -16697.304
Iteration 4: log pseudolikelihood = -16697.304
```

Probit regression

Number of obs = 54,582
Wald chi2(6) = 14383.19
Prob > chi2 = 0.0000
Pseudo R2 = 0.4203

Log pseudolikelihood = -16697.304

ocupinf	Robust					[95% conf. interval]	
	Coefficient	std. err.	z	P> z			
edu	-.137398	.0022802	-60.26	0.000	-.1418671	-.1329289	
exp	-.0479637	.002089	-22.96	0.000	-.0520581	-.0438693	
exp2	.0005228	.0000404	12.94	0.000	.0004436	.000602	
sexo	.1152075	.015596	7.39	0.000	.0846398	.1457751	
nwemp	-.0000282	3.66e-06	-7.70	0.000	-.0000354	-.000021	
tmemp	-.4752976	.0052595	-90.37	0.000	-.4856061	-.4649891	
_cons	3.840976	.0425997	90.16	0.000	3.757482	3.92447	

Fuente: Elaboración propia en base a datos de ENAHO-2022

ANEXO E. Resultado de estimación efectos marginales modelo probit

Marginal effects after probit

y = Pr(ocupinf) (predict)
= .86150975

variable	dy/dx	Std. err.	z	P> z	[95% C.I.]	X
edu	-.0303568	.00046	-65.75	0.000	-.031262 -.029452	9.71146
exp	-.0105971	.00045	-23.30	0.000	-.011488 -.009706	25.9618
exp2	.0001155	.00001	13.10	0.000	.000098 .000133	918.929
sexo	.025454	.00342	7.45	0.000	.018754 .032154	1.47197
nwemp	-6.23e-06	.00000	-7.72	0.000	-7.8e-06 -4.6e-06	1305.63
tmemp	-.1050125	.00155	-67.78	0.000	-.108049 -.101976	1.65677

Fuente: Elaboración propia en base a datos de ENAHO-2022

ANEXO F. Resultados de estimación por método de Oaxaca - Blinder

. oaxaca ingm edu exp exp2 nwemp ttemp, by(ocupinf)

Blinder-Oaxaca decomposition Number of obs = 24,296
 Model = linear
 Group 1: ocupinf = 0 N of obs 1 = 9,266
 Group 2: ocupinf = 1 N of obs 2 = 15,030

endowments: $(X1 - X2) * b2$
 coefficients: $X2 * (b1 - b2)$
 interaction: $(X1 - X2) * (b1 - b2)$

ingm	Coefficient	Std. err.	z	P> z	[95% conf. interval]	
overall						
group_1	7.594967	.0057118	1329.70	0.000	7.583772	7.606162
group_2	6.624896	.0068179	971.70	0.000	6.611533	6.638259
difference	.970071	.0088943	109.07	0.000	.9526386	.9875035
endowments	.5979252	.0142582	41.94	0.000	.5699797	.6258707
coefficients	.4855345	.0120359	40.34	0.000	.4619446	.5091244
interaction	-.1133887	.0163624	-6.93	0.000	-.1454584	-.0813189
endowments						
edu	.2379278	.0077775	30.59	0.000	.2226841	.2531715
exp	.111879	.0079777	14.02	0.000	.0962431	.127515
exp2	-.0302949	.0049441	-6.13	0.000	-.0399852	-.0206046
nwemp	.0011344	.0005924	1.92	0.055	-.0000266	.0022955
ttemp	.2772789	.0125861	22.03	0.000	.2526106	.3019471
coefficients						
edu	.1246232	.0220146	5.66	0.000	.0814755	.167771
exp	-.3715407	.0432805	-8.58	0.000	-.4563689	-.2867125
exp2	.1878036	.0267981	7.01	0.000	.1352802	.240327
nwemp	-.0138118	.0048861	-2.83	0.005	-.0233884	-.0042351
ttemp	-.0796574	.0093272	-8.54	0.000	-.0979384	-.0613763
_cons	.6381174	.037042	17.23	0.000	.5655166	.7107183
interaction						
edu	.0526937	.00933	5.65	0.000	.0344072	.0709802
exp	-.0558622	.0072734	-7.68	0.000	-.0701178	-.0416067
exp2	.0177189	.003696	4.79	0.000	.0104749	.0249629
nwemp	-.0019727	.00076	-2.60	0.009	-.0034622	-.0004832
ttemp	-.1259663	.0147579	-8.54	0.000	-.1548914	-.0970413

Fuente: Elaboración propia en base a datos de ENAHO-2022



ANEXO G. Resultados de estimación método de Ñopo

```
. nopomatch edu exp exp2 nwemp ttemp, outcome(ingm) by(ocupinf)
*****
*****   Gap in edu exp exp2 nwemp ttemp decomposition
*****
D   =-.12772551
D0  =-.04699268
DM  =-.02657566
DF  =-.01856749
DX  =-.03558968
*****
percM =.31949434
percF =.31405137
*****
```

Fuente: Elaboración propia en base a datos de ENAHO-2022



DECLARACIÓN JURADA DE AUTENTICIDAD DE TESIS

Por el presente documento, Yo Pedro Fernando Flores Tacuri
identificado con DNI 72000422 en mi condición de egresado de:

Escuela Profesional, Programa de Segunda Especialidad, Programa de Maestría o Doctorado
Ingeniería Económica.

informo que he elaborado el/la Tesis o Trabajo de Investigación denominada:

"Influencia de la educación y la creación de nuevas empresas
sobre el empleo informal y diferencia de ingresos entre ocupaciones
informales y formales en Peru, periodo 2022"

Es un tema original.

Declaro que el presente trabajo de tesis es elaborado por mi persona y **no existe plagio/copia** de ninguna naturaleza, en especial de otro documento de investigación (tesis, revista, texto, congreso, o similar) presentado por persona natural o jurídica alguna ante instituciones académicas, profesionales, de investigación o similares, en el país o en el extranjero.

Dejo constancia que las citas de otros autores han sido debidamente identificadas en el trabajo de investigación, por lo que no asumiré como tuyas las opiniones vertidas por terceros, ya sea de fuentes encontradas en medios escritos, digitales o Internet.

Asimismo, ratifico que soy plenamente consciente de todo el contenido de la tesis y asumo la responsabilidad de cualquier error u omisión en el documento, así como de las connotaciones éticas y legales involucradas.

En caso de incumplimiento de esta declaración, me someto a las disposiciones legales vigentes y a las sanciones correspondientes de igual forma me someto a las sanciones establecidas en las Directivas y otras normas internas, así como las que me alcancen del Código Civil y Normas Legales conexas por el incumplimiento del presente compromiso

Puno 15 de mayo del 2024

FIRMA (obligatoria)



Huella



AUTORIZACIÓN PARA EL DEPÓSITO DE TESIS O TRABAJO DE INVESTIGACIÓN EN EL REPOSITORIO INSTITUCIONAL

Por el presente documento, Yo Pedro Fernando Flores Tacuri
identificado con DNI 72000422 en mi condición de egresado de:

Escuela Profesional, Programa de Segunda Especialidad, Programa de Maestría o Doctorado

Ingeniería Económica
informo que he elaborado el/la Tesis o Trabajo de Investigación denominada:

“Influencia de la educación y la creación de nuevas empresas sobre el empleo informal y diferencia de ingresos entre ocupaciones informales y formales en Perú, periodo 2022”

para la obtención de Grado, Título Profesional o Segunda Especialidad.

Por medio del presente documento, afirmo y garantizo ser el legítimo, único y exclusivo titular de todos los derechos de propiedad intelectual sobre los documentos arriba mencionados, las obras, los contenidos, los productos y/o las creaciones en general (en adelante, los “Contenidos”) que serán incluidos en el repositorio institucional de la Universidad Nacional del Altiplano de Puno.

También, doy seguridad de que los contenidos entregados se encuentran libres de toda contraseña, restricción o medida tecnológica de protección, con la finalidad de permitir que se puedan leer, descargar, reproducir, distribuir, imprimir, buscar y enlazar los textos completos, sin limitación alguna.

Autorizo a la Universidad Nacional del Altiplano de Puno a publicar los Contenidos en el Repositorio Institucional y, en consecuencia, en el Repositorio Nacional Digital de Ciencia, Tecnología e Innovación de Acceso Abierto, sobre la base de lo establecido en la Ley N° 30035, sus normas reglamentarias, modificatorias, sustitutorias y conexas, y de acuerdo con las políticas de acceso abierto que la Universidad aplique en relación con sus Repositorios Institucionales. Autorizo expresamente toda consulta y uso de los Contenidos, por parte de cualquier persona, por el tiempo de duración de los derechos patrimoniales de autor y derechos conexos, a título gratuito y a nivel mundial.

En consecuencia, la Universidad tendrá la posibilidad de divulgar y difundir los Contenidos, de manera total o parcial, sin limitación alguna y sin derecho a pago de contraprestación, remuneración ni regalía alguna a favor mío; en los medios, canales y plataformas que la Universidad y/o el Estado de la República del Perú determinen, a nivel mundial, sin restricción geográfica alguna y de manera indefinida, pudiendo crear y/o extraer los metadatos sobre los Contenidos, e incluir los Contenidos en los índices y buscadores que estimen necesarios para promover su difusión.

Autorizo que los Contenidos sean puestos a disposición del público a través de la siguiente licencia:

Creative Commons Reconocimiento-NoComercial-CompartirIgual 4.0 Internacional. Para ver una copia de esta licencia, visita: <https://creativecommons.org/licenses/by-nc-sa/4.0/>

En señal de conformidad, suscribo el presente documento.

Puno 15 de Mayo del 2024


FIRMA (obligatoria)



Huella