

**PONTIFICIA UNIVERSIDAD
CATÓLICA DEL PERÚ**

Escuela de Posgrado



Brecha de uso en internet en el Perú. Una descomposición para
áreas rurales

Tesis para obtener el grado académico de Maestro en Regulación de los
Servicios Públicos que presenta:

Marco Antonio Vilchez Román

Asesor:

Gonzalo Martín Ruiz Díaz

Lima, 2025

Informe de Similitud

Yo, Gonzalo Martin Ruiz Diaz, docente de la Escuela de Posgrado de la Pontificia Universidad Católica del Perú, asesor(a) de la tesis/el trabajo de investigación titulado **“Brecha de Uso en Internet en el Perú. Una Descomposición para Areas Rurales”** del autor


MARCO VILCHEZ ROMAN

dejo constancia de lo siguiente:

- El mencionado documento tiene un índice de puntuación de similitud de 12%. Así lo consigna el reporte de similitud emitido por el software *Turnitin* el 23/10/2025.
- He revisado con detalle dicho reporte y la Tesis o Trabajo de Suficiencia Profesional, y no se advierte indicios de plagio.
- Las citas a otros autores y sus respectivas referencias cumplen con las pautas académicas.

Lugar y fecha:

Lima, 07 de Noviembre de 2025

Apellidos y nombres del asesor / de la asesora: Gonzalo Ruiz Diaz	
DNI: 43442771	
ORCID: 0000-0001-9972-9181	
Firma:	

Dedicado:

*A **Lisbell**, mi compañera de vida, por su amor incondicional, su impulso constante y su fe en mí en todo momento. Sin su apoyo, esta meta no habría sido posible.*

*A **José Pablo** y **Carla Sofía**, mis hijos, fuente de inspiración y alegría. Este logro también es para ustedes.*

*A **Manuel** y **Dominga**, mis padres, por enseñarme con su ejemplo el valor del esfuerzo, la perseverancia y la integridad.*

AGRADECIMIENTOS

Agradezco sinceramente a todas las personas que con sus comentarios, consejos y apoyo, contribuyeron con la realización de esta tesis.

De manera especial, quiero mencionar a Gonzalo Ruiz Díaz, mi asesor, por ayudarme a reenfocar mi investigación, así como por su guía, correcciones y recomendaciones, puntuales y oportunas. A Paulo Chahuara Vargas, quien de manera desinteresada y entusiasta me brindó su apoyo para el entendimiento y aplicación de diversos aspectos metodológicos desarrollados. A mi hermano Carlos Vilchez Román por sus observaciones y sugerencias a versiones preliminares de la tesis.

Finalmente, pero siento que de manera primordial, a mi esposa Lisbell Salazar Cáceres, quien desde mucho tiempo atrás estuvo motivándome, e incluso insistiéndome, para cerrar “este ciclo”, y me brindó todo su apoyo incondicional -además de su amorosa comprensión por todo el tiempo dedicado- para la culminación de mi tesis.

Este logro no habría sido posible sin ustedes, gracias.

RESUMEN

Diversa literatura ha evidenciado ya los beneficios para la economía de contar con mayor acceso al Internet, pero el supuesto detrás es que la población utilice efectivamente el referido servicio. Si bien existen iniciativas públicas para fomentar el uso del Internet entre la población (sobre todo más vulnerable), los resultados de encuestas especializadas (ENAHO, ERESTEL) indican que subsiste una brecha urbano-rural importante respecto al uso del Internet. Asimismo, a nivel académico, existe prácticamente un vacío sobre el entendimiento de las brechas de uso del internet en el Perú. Por ello, este documento contribuye primariamente a atender este déficit.

A través del empleo de métodos contrafactuales paramétricos (descomposición de Oaxaca-Blinder) se identifica y cuantifica las características poblacionales observables y no observables que afectan el uso del Internet y que por tanto, brindan luces para el entendimiento y explicación de la brecha o diferencia en el uso del Internet entre regiones urbanas y rurales.

Los resultados obtenidos evidencian que una parte significativa de la referida brecha de uso no se explica por las características observadas de la población y más bien se asocia con factores estructurales (no observables) por el lado de la demanda, como serían las normas sociales, las percepciones de la población rural e incluso limitaciones institucionales. Por ello, la principal recomendación es complementar las medidas por lado de la oferta (mayor infraestructura) con medidas por el lado de la demanda que sean priorizadas, cohesionadas, estructuradas y sostenidas en el tiempo, tales como los programas de alfabetización digital contextualizados y culturalmente sensibles, esquema de subsidios a dispositivos que permitan el uso efectivo.

PALABRAS CLAVE: Brecha de uso de Internet, Descomposición Oaxaca-Blinder, Diferencias Urbano-Rural

ABSTRACT

The economic benefits of increased internet access has already demonstrated by economic literature, and in this regard, on main assumption is that the population actually use the service. Although there are public initiatives to promote internet use among the population, data from specialized peruvian surveys (ENAH0, ERESTEL) indicate that a significant urban-rural gap persists regarding internet use. Besides that, from academic perspective, there is virtually a vacuum regarding the understanding of the gaps in internet use in Peru. Therefore, one main objective of this document is to contribute to addressing this deficit.

Through the use of parametric counterfactual methods (Oaxaca-Blinder decomposition), observable and unobservable population characteristics that affect Internet use are identified and quantified, thus providing insight into and explaining the gap or difference in Internet use between urban and rural regions.

The results obtained show that a significant portion of the aforementioned usage gap cannot be explained by the observed characteristics of the population and is rather associated with structural (unobservable) factors on the demand side, such as social norms, perceptions of the rural population, and even institutional limitations. Therefore, the main recommendation is to complement supply-side measures (better infrastructure) with demand-side measures that are prioritized, cohesive, structured, and sustained over time, such as contextualized and culturally sensitive digital literacy programs and subsidy schemes for devices that enable effective use.

KEY WORDS: Internet usage gap, Oaxaca-Blinder decomposition, Urban-rural

INDICE

ÍNDICE DE TABLAS	8
ÍNDICE DE FIGURAS	9
INTRODUCCIÓN	10
PRIMERA PARTE. MARCO DE LA INVESTIGACIÓN	12
CAPÍTULO 1: DISEÑO DE LA INVESTIGACIÓN.....	12
1.1. EL PROBLEMA	12
1.2. JUSTIFICACIÓN Y OBJETIVOS.....	31
1.3. HIPÓTESIS	33
CAPÍTULO 2: MARCO CONCEPTUAL Y CONTEXTO.....	35
2.1. REVISIÓN DE LA LITERATURA.....	35
2.2. LA BRECHA DE USO DEL SERVICIO DE INTERNET: CARACTERIZACIÓN	40
2.3. DETERMINANTES DE LA BRECHA DE USO DE INTERNET. EL MODELO.....	55
SEGUNDA PARTE. DISEÑO METODOLÓGICO Y RESULTADOS	61
CAPÍTULO 3: PLANTEAMIENTO METODOLÓGICO.....	61
3.1. METODOLOGÍA.....	61
3.2. DESCRIPCIÓN DE VARIABLES Y DATOS	70
3.3. LIMITACIONES	85
CAPÍTULO 4: RESULTADOS Y DISCUSIÓN.....	88
4.1. Estimaciones y resultados.....	88
4.2. DISCUSIÓN DE RESULTADOS Y POSIBLES EXTENSIONES.....	106
CONCLUSIONES.....	115
RECOMENDACIONES	119
REFERENCIAS BIBLIOGRÁFICAS.....	123
ANEXOS.....	135
ANEXO 1. IMPACTO DE LA BANDA ANCHA EN LA ECONOMÍA: RESUMEN	135
ANEXO 2. EFECTOS ESPERADOS DE VARIABLES RELEVANTES EN EL USO DEL INTERNET	136
ANEXO 3. DETALLES DE LAS REGRESIONES	146

ÍNDICE DE TABLAS

Tabla 1 <i>Centros Poblados y población nacional, 2024</i>	18
Tabla 2 <i>Resumen de la literatura</i>	39
Tabla 3 <i>Variables utilizadas en las regresiones</i>	85
Tabla 4 <i>Clasificación de las regiones</i>	89
Tabla 5 <i>Descomposición Oaxaca-Blinder: output del Stata</i>	92
Tabla 6 <i>Primer Modelo. Resultados Descomposición Oaxaca-Blinder</i>	95
Tabla 7 <i>Segundo Modelo. Resultados Descomposición Oaxaca-Blinder</i>	97
Tabla 8 <i>Tercer Modelo. Resultados Descomposición Oaxaca-Blinder</i>	99
Tabla 9 <i>Cuarto Modelo. Resultados Descomposición Oaxaca-Blinder</i>	101
Tabla 10 <i>Quinto Modelo. Resultados de las regresiones por grupos</i>	103
Tabla 11 <i>Quinto Modelo. Resultados Descomposición Oaxaca-Blinder</i>	105
Tabla 12 <i>Resumen de la Descomposición Oaxaca-Blinder para los Modelos 1, 2 y 3</i>	108
Tabla 13 <i>Resumen de la Descomposición Oaxaca-Blinder para los Modelos 4 y 5</i>	109



ÍNDICE DE FIGURAS

Figura 1 Hogares con acceso a internet, 2019-2023	16
Figura 2 Evolución de Sectores en redes móviles: 2015-2024.....	17
Figura 3 Cobertura del servicio móvil: centros poblados y población, 2024.....	19
Figura 4 Cobertura del servicio de internet móvil (4G) e internet fijo: centros poblados y población, 2024	20
Figura 5 Líneas móviles e Internet Móvil, 2015- 2024(Millones)	21
Figura 6 Líneas Internet Móvil y Tráfico promedio, 2015- 2024	21
Figura 7 Conexiones de Internet fijo, 2015-2024 (millones).....	22
Figura 8 Participación de mercado de Internet fijo, 2015-2024.....	23
Figura 9 Conexiones de Internet fijo por tecnología, 2005-2024.....	24
Figura 10 Tarifas promedio de internet fijo e internet móvil 2014-2024 (Sin IGV).....	25
Figura 11 Conexiones de internet fijo por velocidad, 2017-2024	25
Figura 12 Percepción de la velocidad contratada en el servicio de internet fijo, 2019-2024.....	27
Figura 13 Cobertura del servicio de internet móvil (4G) a nivel regional: por segmento urbano-rural y tecnología, 2024.....	41
Figura 14 Cobertura del servicio de Internet por segmento urbano-rural, 2015-2023.....	42
Figura 15 Acceso a internet según tipo de conexión y según ámbito geográfico, 2014-2023....	43
Figura 16 Acceso a internet fijo por regiones, 2023	44
Figura 17 Acceso a internet móvil por regiones, 2023.....	45
Figura 18 Equipamiento TIC del hogar, para uso de Internet: 2019-2023.....	46
Figura 19 Acceso a smartphones según ámbito geográfico, 2019-2023.....	47
Figura 20 Percepción sobre la cantidad de datos del servicio de internet móvil desde el celular, 2019-2023	48
Figura 21 Percepción sobre la velocidad del servicio de internet móvil desde el celular, 2019-2023.....	48
Figura 22 Usos de internet según ámbito geográfico, 2014-2023.....	49
Figura 23 Uso de internet según edad, 2019-2023	50
Figura 24 Uso del internet por sexo, 2014-2023.....	51
Figura 25 Uso del internet por nivel educativo alcanzado como % de la población, 2014 - 2023	52
Figura 26 Formas de Uso del internet según materia de interés, 2014-2023.....	53
Figura 27 Razones para no usar el Internet, 2021-2023	54
Figura 28 Variables que inciden en el Uso de Internet	56
Figura 29 Posible sesgo de agregación de variables.....	87

INTRODUCCIÓN

Una de los objetivos de política en el sector telecomunicaciones en el mundo, ha sido y es la expansión de la infraestructura de telecomunicaciones hacia zonas alejadas de las grandes ciudades (denominada también expansión de la cobertura), y el Perú no ha sido ajeno a ello. Las políticas de universalidad han sido las herramientas para el logro de tal fin, surgiendo conceptos como acceso universal, servicio universal, cobertura universal geográfica, frontera de eficiencia de mercado, frontera de sostenibilidad, brecha de eficiencia de mercado, brecha de acceso, los cuales son ampliamente explicados por Stern y Townsend (2007) y por Navas-Sabater et al. (2002).

Si bien se ha dado una importante expansión del servicio de acceso a internet, sobre todo móvil, aún quedan espacios por ser atendidos, y no nos referimos únicamente a la dimensión geográfica.

Diversa literatura ha evidenciado ya los beneficios para la economía de contar con mayor acceso al internet. Dichos efectos se verían potenciados aun más si las personas logran apropiarse del internet, de su contenido, y para ello el uso efectivo del internet se hace indispensable.

En el Perú se han desplegado algunas iniciativas, desde diversas aristas del ordenamiento público, orientadas a fomentar el uso del internet; sin embargo, hasta el momento no se ha logrado emparejar u homogenizar tal uso a nivel nacional. ¿Tendría que serlo? ¿dichas intervenciones deben ser homogéneas, o diferenciadas según regiones o según colectivos socio-económicos?

A nivel académico, existe prácticamente un vacío sobre el entendimiento de las brechas de uso del internet en el Perú. Por ello, este documento contribuye primariamente a atender este déficit.



PRIMERA PARTE. MARCO DE LA INVESTIGACIÓN

CAPÍTULO 1: DISEÑO DE LA INVESTIGACIÓN

1.1. EL PROBLEMA

Una de los objetivos de política en el sector telecomunicaciones en el mundo, ha sido y es la expansión de la infraestructura de telecomunicaciones hacia zonas alejadas de las grandes ciudades (denominada también expansión de la cobertura), y el Perú tampoco ha sido ajeno a ello. Las políticas de universalidad han sido las herramientas para el logro de tal fin, surgiendo conceptos como acceso universal, servicio universal, cobertura universal geográfica, frontera de eficiencia de mercado, frontera de sostenibilidad, brecha de eficiencia de mercado, brecha de acceso, los cuales son ampliamente explicados por Stern y Townsend (2007) y por Navas-Sabater et al. (2002):

- Brecha de acceso: según Stern y Townsend (2007) la brecha de acceso es la diferencia entre la frontera de eficiencia del mercado y un nivel de penetración nacional del 100 por ciento¹. Este concepto hace referencia al déficit en la disponibilidad de la infraestructura de telecomunicaciones y/o de dispositivos para la población.
- Brecha de uso: La brecha de uso tiene que ver con el déficit en el uso del servicio de Internet, aun cuando se cuente con acceso a la red de telecomunicaciones y/o a los

¹ De acuerdo con el autor, la frontera de eficiencia del mercado hace referencia al nivel de penetración que naturalmente puede ser alcanzada en un mercado competitivo que se caracteriza por un adecuado desempeño y que cuenta con un entorno regulatorio estable. Esto supone que existe “algún tamaño teórico de mercado “eficiente” para cualquier servicio dado, que los operadores competitivos y con fines de lucro aceptarían proveer sin ningún subsidio ni otras ayudas externas, siempre y cuando se eliminen las barreras artificiales (no económicas) a la entrada y a la expansión. La ubicación de esta frontera de eficiencia del mercado está principalmente determinada por factores del lado de oferta.” (Stern 2007, pg 57)

dispositivos que permiten tal uso. Por ello, también se hace referencia a las capacidades que se tienen o que se necesitan para hacer un uso efectivo de las TIC para así apropiarse de su contenido y beneficios derivados. Las barreras a las que se enfrentan las poblaciones incluyen el analfabetismo digital o falta de habilidades digitales, asequibilidad (especialmente para adquirir equipos móviles o computadoras/laptops), acceso a contenidos y servicios relevantes, preocupaciones sobre seguridad y acceso, entre otras.

- Brecha digital²: se refiere a la brecha entre individuos, hogares, empresas, áreas geográficas de diferentes niveles socioeconómicos en cuanto a las oportunidades de acceder (conectividad a internet y acceso a los dispositivos habilitados para usar Internet) y usar las tecnologías de la información y comunicación (TIC) especialmente los servicios basados en el Internet, para una gama amplia de actividades. Dichos aspectos han sido analizados por la Organisation for Economic Co-operation and Development [OECD] (2001), OECD (2002) y Park et al. (2015). La brecha digital también incluye aspectos referidos a las habilidades de alfabetización digital (capacidad para utilizar adecuadamente las TIC) debido a los diferentes niveles de educación y conocimientos técnicos (Naciones Unidas, 2021; Hincu et al., 2011 y CEPLAN³).

En esa línea, la conectividad a internet es la base para la participación en la sociedad digital: sin acceso a internet, eso es imposible. Por ello, se debe buscar una

² Un concepto similar es el de equidad digital, sobre el cual la Internet Society Foundation (<https://www.isocfoundation.org/2023/06/what-is-digital-equity/>, búsqueda realizada el 20 de abril de 2025) señala que se refiere a la situación en la que cada persona y cada comunidad tiene los recursos de tecnología de la información necesarios para participar adecuada y completamente en la sociedad y la economía. Por tanto, el concepto comprende tanto el acceso a dispositivos y al internet, sino también la capacidad de darles un uso eficaz para poder beneficiarse de ellos de manera efectiva.

³ Ver (https://observatorio.ceplan.gob.pe/ficha/o9_lamb), búsqueda realizada el 24 de enero de 2025.

conectividad a internet que sea “robusta, sostenible, asequible e inclusiva”, pues de lo contrario, los beneficios de las tecnologías digitales, la participación en la sociedad digital y el acceso a la oferta de servicios digitales, serán sistemáticamente excluyentes (Naciones Unidas, 2021).

Existe diversa literatura que evidencia los beneficios para la economía de contar con mayor acceso al internet⁴, sobre todo por su efecto en la producción nacional, en el crecimiento de los países (en el Anexo 1 se presenta un breve resumen de los mismos), e incluso en la productividad nacional. Así, la ITU (2021-i) señala que:

- A nivel mundial agregado, un incremento de 10% en la penetración de la banda ancha genera un incremento en la producción nacional de entre 0.77%-0.80% en el caso de la banda ancha fija y entre 1.5%-1.6% para la banda ancha móvil.
- En países desarrollados el efecto es mayor: un incremento de 10% en la penetración de la banda ancha fija genera un incremento en la producción nacional per cápita de 1.25%, mientras que en países de mediano ingreso el efecto es 0.85% y en países de bajo ingreso, los efectos son no significativos. En el caso de la banda ancha móvil el efecto es de 2.04% en el Producto Bruto Interno (PBI) per cápita para países con menores ingresos, mientras que en los países de ingresos medianos el efecto es 1.62%.

⁴ Conviene aclarar una precisión respecto a la delimitación de la revisión de la literatura. El presente trabajo toma como supuesto o punto de partida que efectivamente el mayor acceso y uso del Internet genera beneficios a la población y a la economía. Es por este motivo que la revisión de literatura o las referencias bibliográficas, en este extremo, consideran únicamente aquellos estudios previos que han cuantificado estos efectos positivos del acceso del Internet sobre la economía. En caso, el objetivo fuera analizar el acceso del Internet, la búsqueda y consideraciones bibliográficas hubieran sido más amplias.

Por su parte, Analysys Mason (2017) resume los principales efectos económicos derivados de un incremento en la penetración de la banda ancha⁵:

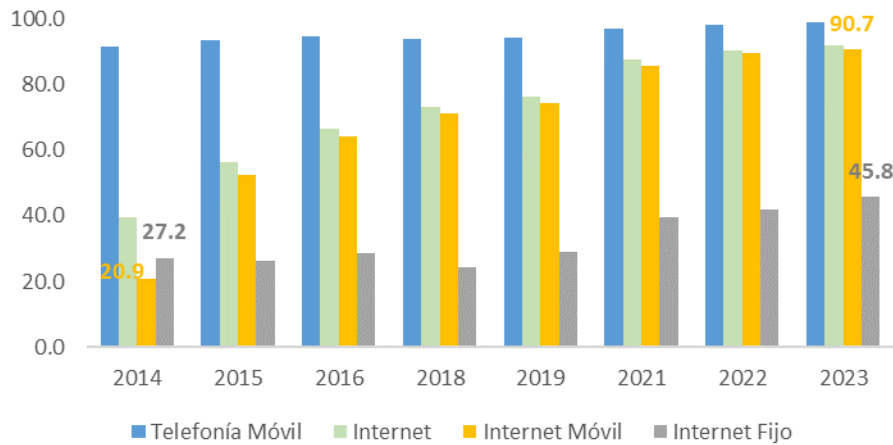
- i. Efectos directos (corto plazo): actividad económica y nuevos empleos directamente asociados con el despliegue de la infraestructura de red.
- ii. Efectos indirectos (corto a mediano plazo): mejoras en la productividad gracias a la adopción de procesos más eficientes (mejoras en términos de tiempo, evitando traslados y viajes).
- iii. Efectos inducidos (mediano a largo plazo): nuevas formas de hacer negocios impulsadas por la innovación, el mayor acceso a la información y la mayor disposición a la tecnología.

En el Perú se ha avanzado con el despliegue del servicio de acceso a internet, evidentemente mucho más a nivel del internet móvil que respecto del internet fijo. La Figura N° 1 muestra la tasa de acceso de hogares a servicios públicos de telecomunicaciones, donde se muestra que al cierre del 2023, casi el 99% de los hogares en el Perú cuenta con acceso al servicio móvil (voz y/o internet) y el 90,7% de hogares cuenta con acceso al servicio de internet móvil; mientras que en el caso de internet fijo, la tasa de acceso llega a casi el 46% de los hogares a nivel nacional.

⁵ Estos efectos son mostrados en la “Figura 01 Principales efectos económicos derivados de un aumento de la penetración de la banda ancha”, en la página 12.

Figura 1

Hogares con acceso a internet, 2019-2023

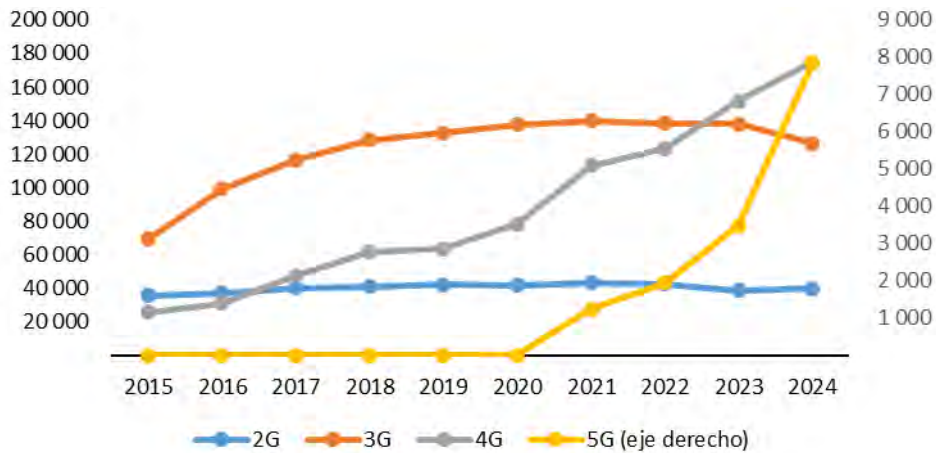


Elaboración propia a partir de la Erestel (OSIPTEL, varios años)

En el presente estudio se asume que el servicio de internet móvil puede ser mejor aprovechado –es decir, que se logra una mejor apropiación de sus contenidos y en consecuencia se pueden obtener los beneficios ya señalados por la literatura de un mayor acceso al internet (de banda ancha)- a partir de la tecnología 4G (y no desde la tecnología 3G). La primera razón considerada para dicho supuesto es que, la calidad de la experiencia con los contenidos que se vienen dando a través del internet se torna adecuada a partir de conexiones 4G (el acceso a internet con redes 2G y 3G, si bien es posible, no permite una adecuada experiencia a los usuarios debido a la lentitud o intermitencias). De otro lado, si se analiza la evolución de los sectores tecnológicos de las redes móviles, se aprecia que los sectores 3G vienen disminuyendo en comparación con los sectores 4G, y se estima que esa tendencia se acreciente en el corto plazo.

Figura 2

Evolución de Sectores en redes móviles: 2015-2024



Elaboración propia a partir de estadísticas OSIPTEL (NRIP al 2024)

En tercer lugar, el Ministerio de Transportes y Comunicaciones (MTC) dentro de la documentación oficial referida al proceso de licitación de las bandas 5G⁶ ha señalado que donde no existe tecnología 4G, “no hay internet móvil”; entonces desde el punto de vista del desarrollador de la política sectorial, internet móvil se asocia a tecnologías 4G o superiores. En cuarto lugar, si se analizan las ofertas comerciales del servicio de internet, se puede ver que los atributos de los nuevos planes de internet fijo (básicamente con fibra óptica) son similares a los planes de internet móvil con redes 4G por lo menos. Y si se

⁶ El Decreto Legislativo N° 1627, Decreto Legislativo para promover el despliegue de los servicios públicos de telecomunicaciones mediante tecnología de quinta generación (5G) o superior (publicado en el Diario Oficial El Peruano el 17 de agosto de 2024); el Decreto Supremo N° 004-2025-MTC (publicado en el Diario Oficial El Peruano el 27 de marzo de 2025) mediante el cual se aprueba el Reglamento del Decreto Legislativo N° 1627; la Resolución Ministerial N° 199-2025-MTC/03 (publicada en el diario oficial El Peruano el 16 de abril de 2025) mediante la cual se aprobó la Convocatoria al Mecanismo Especial de Asignación y acondicionamiento de la banda de frecuencias de 3.5 GHz (3 300 - 3 800 MHz) para la prestación de servicios públicos de telecomunicaciones mediante tecnología de quinta generación (5G) o superior. Nótese que, tal como lo indica la Resolución Ministerial N° 199-2025-MTC/03 “con la Convocatoria se establecieron, entre otros, los requisitos legales, técnicos y económico – financieros que deben cumplir todas las empresas interesadas para acceder al mecanismo de asignación especial y los requisitos que deben cumplir los concesionarios que opten por el mecanismo de Acondicionamiento; así como, el Cronograma para la conducción del Mecanismo especial de asignación y Acondicionamiento, estableciendo plazos y etapas para realizar diversas actuaciones”.

comparan las velocidades de internet móvil de planes 3G versus los planes 4G, la diferencia es notable.

Cobertura

Considerando la población reportada en el Censo Nacional 2017, el 99% de la población peruana se concentra o vive en el 52% de los centros poblados a nivel nacional (en el Perú existen 108 228 centros poblados, de los cuales 104 464 son centros poblados rurales). En el Perú hay 86 961 centros poblados rurales (con menos de 100 habitantes), y en dichos centros poblados hay una población de 1.63 millones de habitantes⁷.

Tabla 1

Centros Poblados y población nacional, 2024

Habitantes	Número de CCPP			Población involucrada		
	Urbano	Rural	Total	Urbano	Rural	Total
0	1	20 367	20 368			
1 - 10	5	31 153	31 158	18	124 786	124 804
11 - 50	12	23 085	23 097	356	604 570	604 926
51 - 100	75	12 264	12 339	5 931	891 945	897 876
101 - 500	1 351	16 701	18 052	419 726	3 314 580	3 734 306
501 - 1000	922	772	1 694	651 851	505 360	1 157 211
> 1000	1 398	122	1 520	23 145 411	180 533	23 325 944
Total	3 764	104 464	108 228	24 223 293	5 621 774	29 845 067

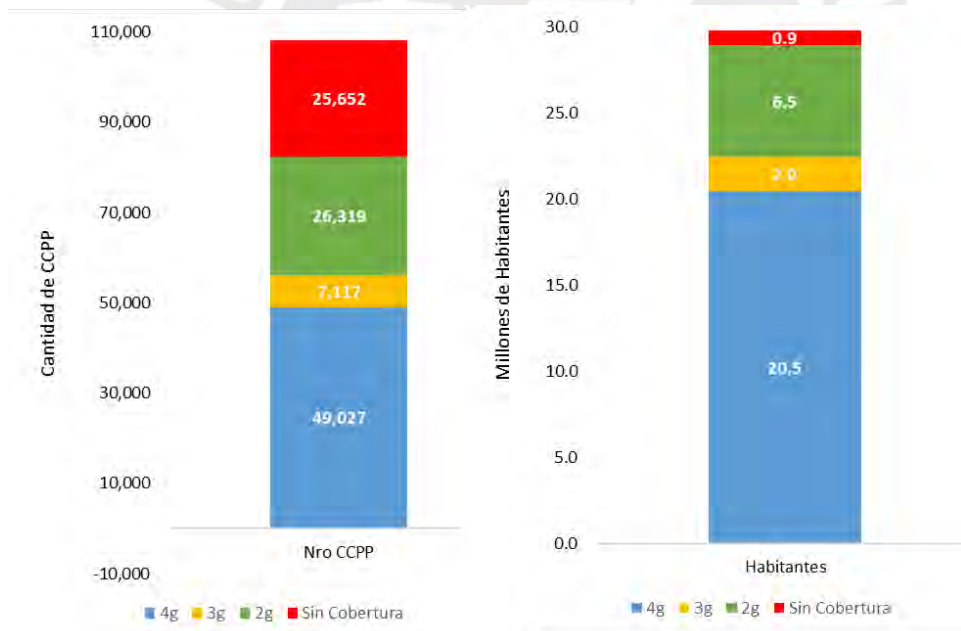
Elaboración propia a partir del Directorio Institucional de Centros Poblados del OSIPTEL (NRIP).

⁷ En el Directorio Institucional de Centros Poblados del OSIPTEL (cuya referencia y actualizaciones dependen de los datos del INEI, y cuya última actualización fue el 2023 a través de la Resolución 176-2023-GG/OSIPTEL del 20 de mayo de 2023), la población total nacional a nivel de centros poblados es de 29 808 780 habitantes y la cantidad total de centros poblados es de 108 115, de los cuales 3 754 son urbanos y 103 959 son rurales (402 centros poblados contenidos en el referido Directorio no se encuentran en la base del INEI). Si bien hay diferencias entre los diversos listados que se manejan a nivel sectorial (como los listados del MTC y del PRONATEL), en general los datos contenidos son similares; y las conclusiones derivadas sobre los datos de dichos listados son similares, también.

El servicio móvil es el que cuenta con mayor despliegue a nivel nacional. Así, al cierre del año 2024, existen 82 463 centros poblados con declaración de cobertura móvil, los cuales concentran el 97% de la población nacional. De este total 78 728 son centros poblados rurales, los cuales congregan a 4.65 millones de habitantes. Asumiendo que la tecnología 4G en adelante permite el uso adecuado y la apropiación de los contenidos del internet, se tiene que 49 027 centros poblados a nivel nacional cuentan con cobertura del internet móvil (concentrando a 20.5 millones de habitantes, que representa el 69% de la población nacional).

Figura 3

Cobertura del servicio móvil: centros poblados y población, 2024

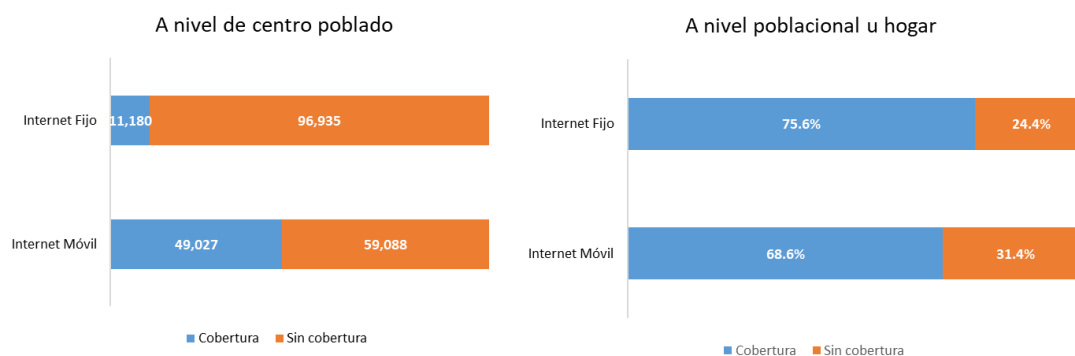


Elaboración propia a partir de estadísticas OSIPTEL (NRIP al 2024)

Por el lado del internet fijo, la cobertura de este servicio abarca 11 180 centros poblados a nivel nacional (que albergan al 75,6% de los hogares a nivel nacional), nivel inferior a los 49 027 centros poblados que cuentan con acceso a internet móvil.

Figura 4

Cobertura del servicio de internet móvil (4G) e internet fijo: centros poblados y población, 2024



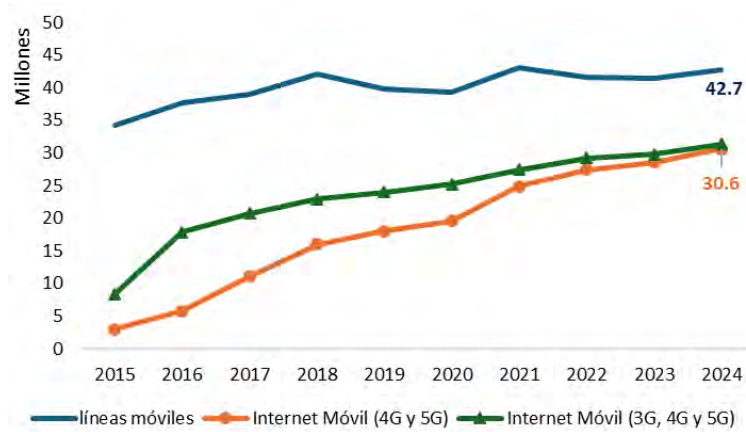
Elaboración propia a partir de estadísticas OSIPTEL (NRIP al 2024)

Líneas o Conexiones

Al cierre de 2024, en el Perú habían 42.7 millones de líneas móviles en servicio, de los cuales 30,6 millones accedieron a internet mediante las tecnologías 4G y 5G (si además se agrega la tecnología 3G, el número de líneas móviles que accedieron a internet asciende a 31,3 millones).

Figura 5

Líneas móviles e Internet Móvil, 2015- 2024(Millones)

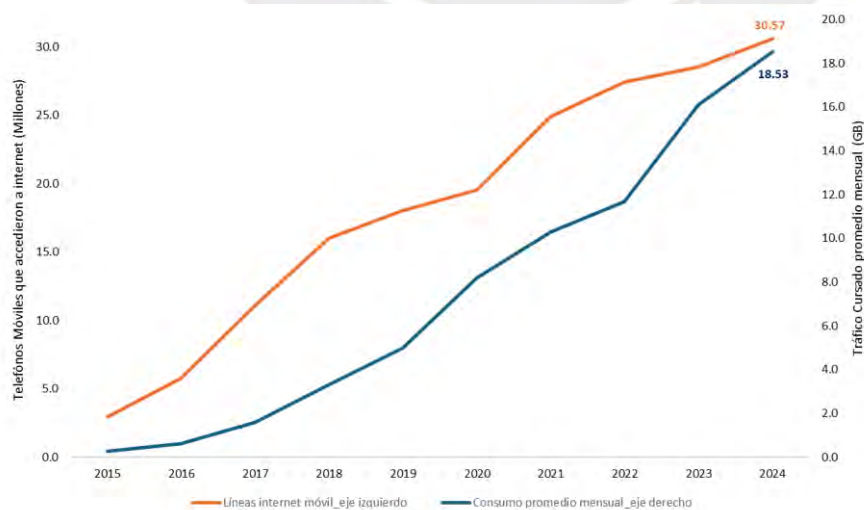


Elaboración propia a partir de estadísticas OSIPTEL (NRIP al 2024)

En los últimos 9 años, el nivel del tráfico de internet móvil promedio mensual por línea se incrementó en más de 67 veces, pasando de 0,27 GB en 2015 a 18,5 GB a fines de 2024; lo cual refleja que el uso del internet móvil ha incrementado sustantivamente a nivel nacional.

Figura 6

Líneas Internet Móvil y Tráfico promedio, 2015- 2024

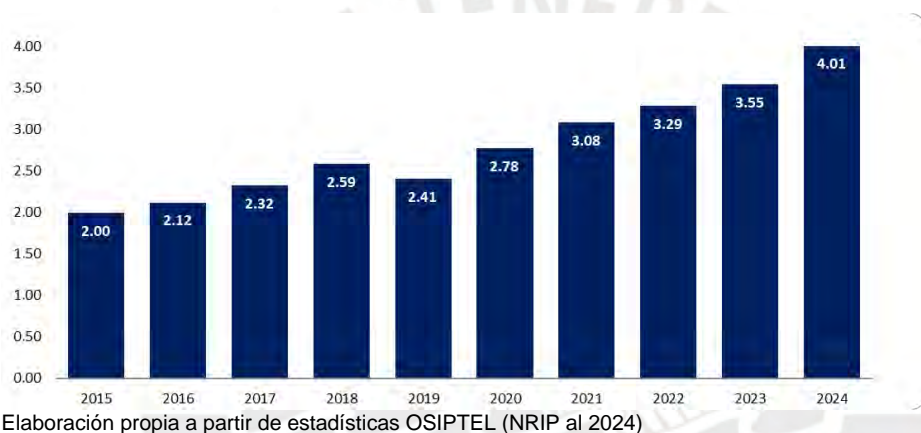


Elaboración propia a partir de estadísticas OSIPTEL (NRIP al 2024)

Por el lado de las conexiones de internet fijo, su evolución creciente de los últimos años se explica básicamente por la entrada de nuevos operadores que vienen desplegando sus redes de fibra óptica. Entre el año 2015 y el 2024, las conexiones de internet fijo pasaron de 2 millones a poco más de 4 millones, es decir se duplicaron en el lapso de 9 años.

Figura 7

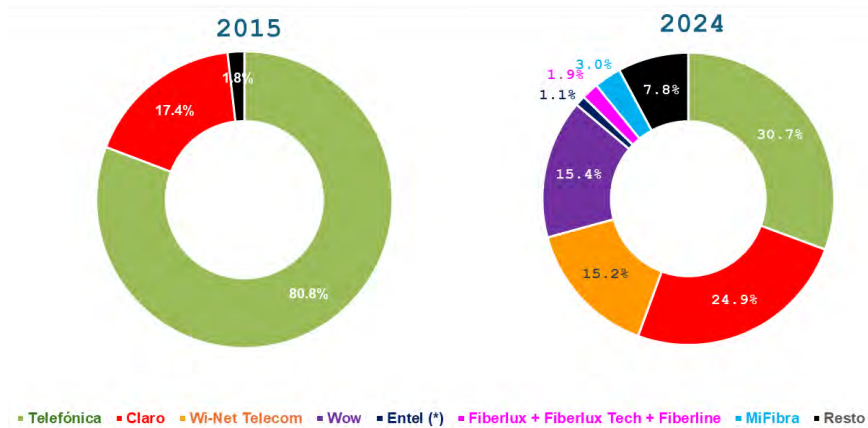
Conexiones de Internet fijo, 2015-2024 (millones)



El nivel de competencia en el mercado de internet fijo se ha incrementado positivamente en los últimos 9 años. De un mercado altamente concentrado donde una sola empresa era la líder indiscutible, se ha pasado a una estructura más equilibrada donde los nuevos competidores importantes han desplegado sus redes con la tecnología de fibra óptica.

Figura 8

Participación de mercado de Internet fijo, 2015-2024

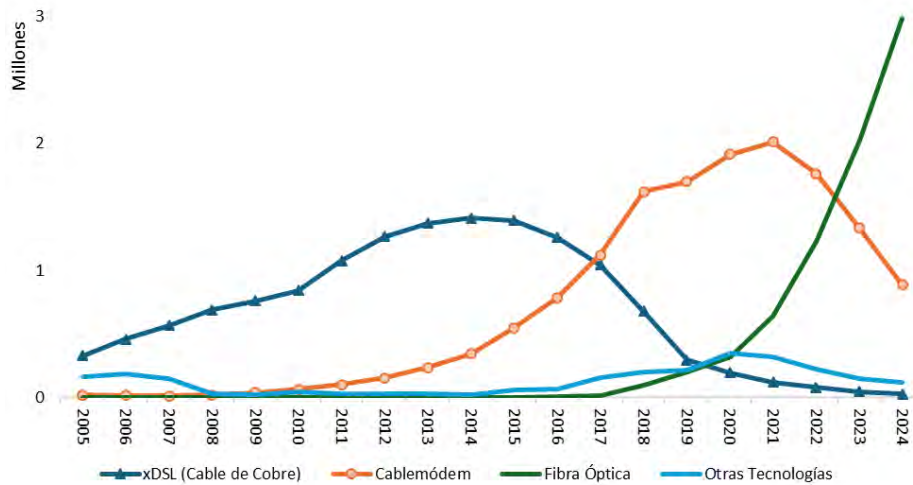


Elaboración propia a partir de estadísticas OSIPTEL (NRIP al 2024)

Los nuevos despliegues de internet fijo se vienen dando a través de la fibra óptica, e incluso las empresas operadoras han iniciado ya planes de renovación tecnológica del referido servicio. Así, la fibra óptica se consolida como el principal medio para acceder a internet fijo. A fines de 2024, las conexiones a internet fijo mediante la fibra óptica (2.9 millones) representaron el 74.5% del total de conexiones de internet fijo del mercado.

Figura 9

Conexiones de Internet fijo por tecnología, 2005-2024

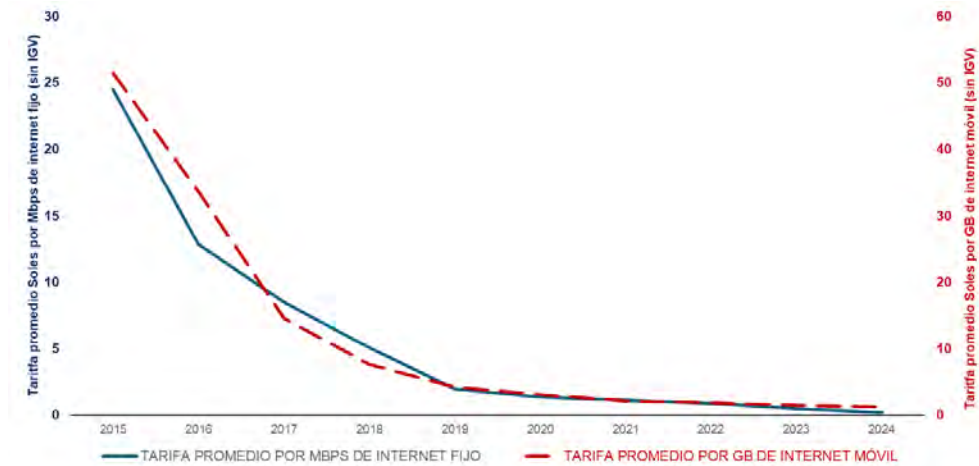


Elaboración propia a partir de estadísticas OSIPTEL (NRIP al 2024)

La dinámica competitiva en ambos mercados (internet móvil e internet fijo) ha estado centrada en ofrecer menores tarifas a los usuarios; y en los últimos años se ha incrementado el ofrecimiento de atributos en los planes tarifarios (mayor velocidad o mayor capacidad). En el gráfico siguiente se puede apreciar que la tarifa promedio de internet fijo (por Mbps) ha se reducido 99% en los últimos 9 años; mientras que la tarifa promedio del internet móvil (por GB) lo ha hecho en 98%.

Figura 10

Tarifas promedio de internet fijo e internet móvil 2014-2024 (Sin IGV)

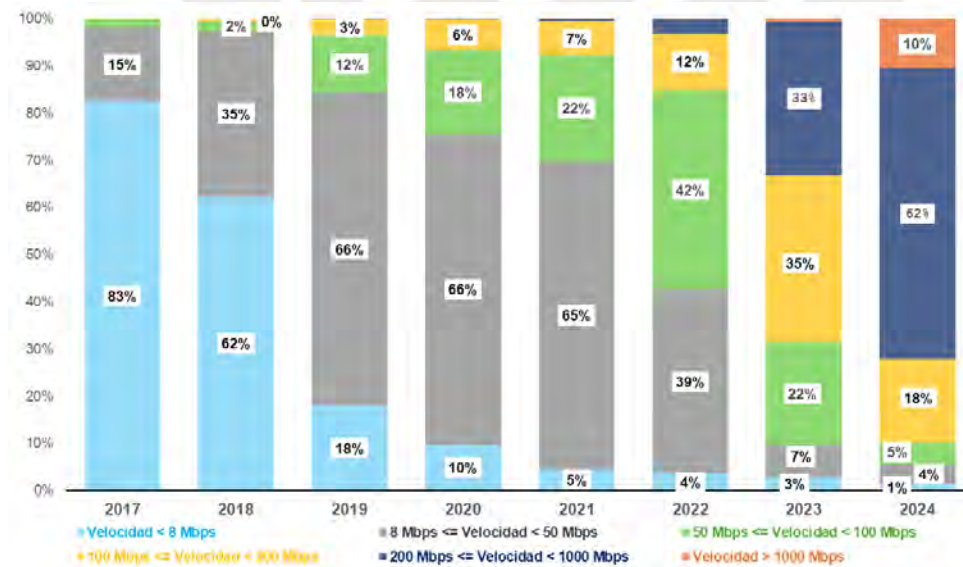


Elaboración propia a partir de estadísticas OSIPTEL (NRIP al 2024)

Al cierre del 2024, el 90% de las conexiones de internet fijo ofrecen velocidades superiores a los 100 Mbps.

Figura 11

Conexiones de internet fijo por velocidad, 2017-2024



Elaboración propia a partir de estadísticas OSIPTEL (NRIP al 2024)

Los resultados obtenidos a partir de la Norma de Requerimientos de Información Periódica (NRIP)⁸ del Organismo Supervisor de Inversión Privada en Telecomunicaciones (OSIPTEL) se condicen con la información contenida en la Encuesta Residencial de Servicios de Telecomunicaciones (ERESTEL) del OSIPTEL. De acuerdo a la referida encuesta del año 2023, y tal como muestra en la Figura 12, casi el 32% de los hogares contrataron el servicio de acceso a internet con altas velocidades de navegación (en el caso de la ERESTEL informaron velocidades mayores a los 60 Mbps), siendo que esta tendencia se ha mantenido creciente en los últimos 4 años.

Los datos mostrados, evidencian una mayor presencia del servicio de internet a nivel nacional, lo cual a su vez es reflejo de una mayor competencia en dichos mercados o segmentos; es decir, la oferta del servicio de acceso a internet (lo que se ofrece) ha mejorado sustantivamente en el Perú.

Si bien se ha dado una importante expansión del servicio de acceso a internet, sobre todo móvil, aún quedan espacios por ser atendidos, y no nos referimos únicamente a la dimensión geográfica.

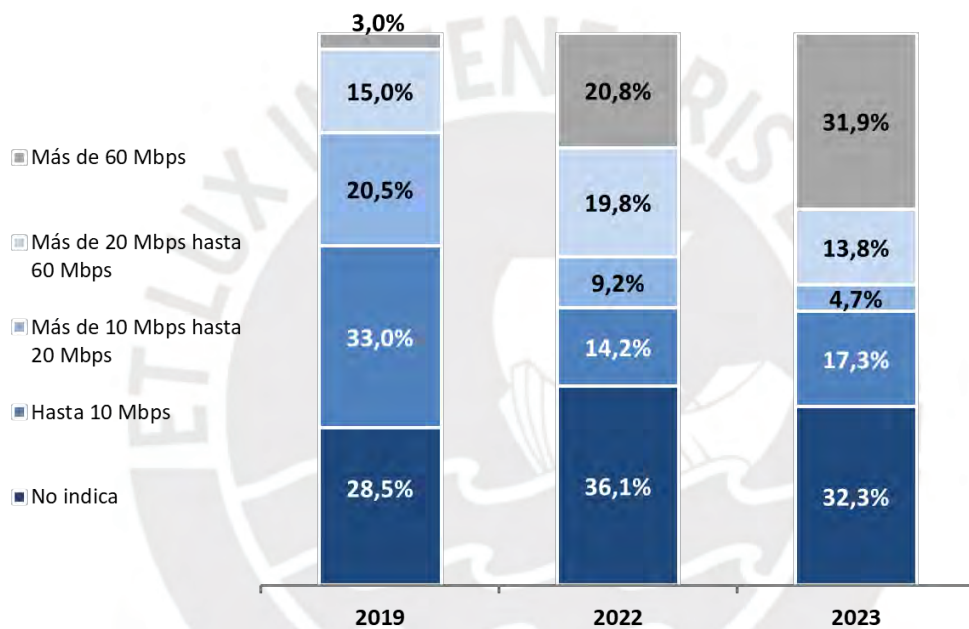
En esa línea, bien cabría preguntarse ¿qué hay del uso del internet en el Perú? ¿cómo se viene dando el uso (o la apropiación) del internet en la población? ¿los efectos de un mayor acceso a internet están llegando a toda la población? La apropiación (el uso) del internet depende del acceso, esto es: sin acceso no hay disfrute del internet, por ello la primera tarea a ser atendida fue (y es) incrementar el acceso a internet.

⁸ Aprobada mediante la Resolución de Consejo Directivo N° 043-2022- CD/OSIPTEL y modificatorias.

Queda claro que dichos efectos de contar con internet, de acceder al internet, se verían potenciados aun más si las personas logran apropiarse del internet, logran beneficiarse directamente de los contenidos (positivos y productivos) del internet, y para ello el uso efectivo del internet se hace indispensable.

Figura 12

Percepción de la velocidad contratada en el servicio de internet fijo, 2019-2024



Elaboración propia a partir de la Erestel (OSIPTEL, varios años)

En su reporte del 2024, la Global System for Mobile Communication Association (GSMA) describe la brecha general de conectividad móvil (que incluye la brecha de acceso y la brecha de uso) a nivel mundial, siendo sus principales hallazgos:

- 4.6 billones de personas (57% de la población mundial) usan el internet móvil desde sus propios dispositivos
- 350 millones de personas (4% de la población mundial) viven en grandes áreas remotas/alejadas que no cuentan con cobertura de redes de internet móvil (ésta es la brecha de cobertura)

- Si bien el 96% de la población mundial vive en zonas que cuentan con cobertura de redes móviles de banda ancha, existe cerca de 3.1 billones de personas a nivel mundial (que equivalen al 39% de la población mundial), quienes *viven en zonas con cobertura móvil de banda ancha⁹, pero enfrentan barreras que les impide mantenerse conectada o usar eficazmente el internet (esta es la brecha de uso)*. La brecha de uso es 9 veces el tamaño de la brecha de cobertura.

Existen diversos estudios que analizan la brecha de acceso del servicio de (acceso a) internet, pero no muchos que hagan lo propio con la brecha de uso y con la brecha digital, al menos para el Perú. Lo que sí hay son estudios que caracterizan la demanda por uso del internet a partir de descripciones de los resultados de diversas encuestas como la Encuesta Nacional de Hogares (ENAH) del Instituto Nacional de Estadísticas e Informática (INEI) y la ERESTEL, tales como los desarrollados por Castillo y Chahuara (2020), Gavilano y Castillo (2019), INEI (2024), Ministerio de Transportes y Comunicaciones [MTC] (2023) y OSIPTEL (2023).

De otro lado, si bien en el Perú se han desplegado algunas iniciativas, desde diversas aristas del ordenamiento público, orientadas a fomentar el uso del internet, está claro que el uso del referido servicio no es parejo u homogéneo a nivel nacional. En primer lugar, se puede mencionar al Plan Nacional de Banda Ancha del año 2011, como la primera iniciativa con alcance nacional, el cual luego de presentar un diagnóstico sobre el desarrollo de la banda ancha en el país y las barreras que limitan ese desarrollo, plantea las metas y propuestas de política para el desarrollo de la banda ancha en el Perú¹⁰.

⁹ La GSMA indica que casi el 90% de la población mundial que no usa el Internet móvil vive en zonas donde sí existe cobertura del referido servicio (banda ancha móvil).

¹⁰ En particular las recomendaciones 10.2.4 (Incentivar la creación y desarrollo de contenidos y aplicaciones digitales en el país) y 10.2.5 (Contar con un plan estratégico de capacitación nacional orientado a desarrollar capacidades y habilidades en el uso de TIC en la población), han permitido el desarrollo de las iniciativas que se mencionan a continuación.

De una lado, el Ministerio de Educación (MINEDU) estableció el 2016 el Plan Nacional de Alfabetización Digital (PNAD) que tiene por objetivo impulsar el uso de la banda ancha en todo el país, promoviendo infraestructura, servicios y habilidades digitales para mejorar la inclusión social, el desarrollo económico, la competitividad y la seguridad. Tiene como meta para el 2030, aumentar sustancialmente el número de jóvenes y adultos que tienen las competencias necesarias, en particular técnicas y profesionales, para acceder al empleo, el trabajo decente y el emprendimiento. Mantiene un eje de coordinación con el MTC, con la Oficina Nacional de Gobierno Electrónico e Informática de la PCM (ONGEI), y con la DITE Unidad de Estadística (UE).

Asimismo, el MINEDU cuenta con otras iniciativas tales como el Programa de Alfabetización (PACE) iniciado el 2022 con alcance en Piura, y que busca cerrar brechas del sector educativo; el Programa Aprendiendo Emprendo (2022) destinado a fortalecer la capacitación de comerciantes en el espacio público mediante el uso de la educación presencial y virtual, para mejorar sus competencias y capacidades en emprendimiento.

Por su parte el MTC estableció el 2023, el Programa Vacaciones Útiles 2023 con el fin de Fortalecer las competencias digitales de niños y adolescentes en zonas rurales, para que puedan hacer un mejor uso de las Tecnologías de la Información y Comunicación (TIC). El ámbito geográfico comprende Apurímac, Ayacucho, Huancavelica, Huánuco, Junín y Pasco. Se brindaron cuatro cursos: empoderamiento digital, Word básico, Excel básico y Power Point. Alcanzando a 1,885 niños y adolescentes capacitados entre el 5 de enero y el 15 de marzo de 2023.

Luego, el Ministerio de Cultura inauguró el 2023 el Programa En la biblioteca aprendo, que busca brindar conocimientos y desarrollar competencias tecnológicas en los participantes, permitiéndoles mejorar en el manejo de computadoras y el uso de

herramientas digitales como Word, Excel, redes sociales, correo electrónico y navegación en línea. Este curso está dirigido principalmente a los adultos mayores.

Por su parte, el Programa Nacional de Telecomunicaciones (PRONATEL) cuenta con una serie de iniciativas como los Centros de Atención Digital (CAD) que desde el 2022 promueven la alfabetización e inclusión digital en zonas rurales; los Espacios Públicos de Acceso Digital (EPAD) que brindan acceso libre y gratuito a Internet vía WiFi en plazas o espacios públicos estratégicos, permitiendo que la población se conecte mediante sus dispositivos personales (smartphones, tabletas, etc.) dentro del área de cobertura.

Al cierre de 2023, se concretaron 9 636 atenciones realizadas en 104 CAD: 8 130 ciudadanos accedieron al servicio de Internet y 1239 realizaron trámites públicos virtuales. Se impartieron 2860 capacitaciones, donde el 52.6% fue realizada por líderes digitales y 44.7% mediante cursos virtuales de aliados Netzum y EDTeam. El 52% de los usuarios fueron mujeres, con una edad promedio de 34 años, y el 22.6% hablaban quechua.

Espacios Públicos de Acceso Digital (EPAD), están próximos a implementar más de 4,780 nuevos espacios digitales en 3,722 centros poblados de 23 regiones de nuestro país. Esta intervención se llevará a cabo a través de proyectos como EPAD Selva Fase 1 y 2, Conecta Móvil+, y EPAD Huancavelica, Cusco, Ayacucho y Apurímac.

Por tanto, el problema que se identifica es que, para el caso peruano, no existe un estudio sobre los determinantes de las brechas de uso del servicio de acceso a internet, sobre los factores que explican las diferencias en el uso del internet entre población urbana y población rural.

Por ello, se plantea una aproximación empírica para analizar las brechas en el uso de internet en el Perú entre la población urbana y la población rural, empleando la técnica

de descomposición Oaxaca-Blinder. De esta manera, se analiza si esta brecha tiene un componente que se explica por factores no observados y cuánto se explica principalmente por características observables, (ie, cuánto de la brecha en el uso del internet se debe o se explica por dichas características), donde:

- a) Características/factores observables: incluyen a la educación, ingresos, infraestructura básica, etc. (son factores que pueden ser controlados por el gobierno),
y
- b) Características/factores no observables: tiene que ver con aspectos relacionados con la cultura digital o alfabetización digital, barreras culturales, fallas de mercado, etc..

La descomposición Oaxaca-Blinder es una técnica desarrollada originalmente para estudiar diferencias salariales entre grupos (como hombres y mujeres), pero se ha extendido a diversos campos donde se busca explicar brechas observadas en una variable de interés entre dos grupos poblacionales. En términos prácticos, la descomposición O-B permite “explicar qué parte de la diferencia entre dos grupos se debe a las diferencias en las variables explicativas y qué parte se debe a las diferencias en los coeficientes de regresión”; de esta manera se puede “identificar la contribución de cada factor en la desigualdad”.

1.2. JUSTIFICACIÓN Y OBJETIVOS

Hay un vacío en la literatura respecto a la identificación de los determinantes del (la tasa de) uso del servicio de (acceso a) internet, para luego poder explicar diferencias entre grupos poblacionales.

Así, las preguntas que se buscarían responder son ¿de qué depende la tasa de uso del internet a nivel regional?, ¿qué parte de la brecha en uso de internet se debe a diferencias en características observables (educación, ingresos, pobreza, etc.), ¿qué parte se mantiene no explicada por factores observables?, ¿la brecha de internet se debe a falta de infraestructura (explicado) o a diferencias en componentes no observados (e.g. fallas de mercado, eficiencia en cómo se usan los recursos de los programas, déficit de alfabetización digital, preferencias o aversiones en el uso de internet)?

Contar con respuestas a las preguntas planteadas, permitiría atender de mejor manera la preocupación de ¿cómo fomentar el uso del internet para incrementar la productividad y el bienestar?.

Si bien existen iniciativas públicas al respecto, ¿se han definido de manera estructurada? ¿existe un plan nacional al respecto?. Conocer, entonces, la naturaleza en la brecha de uso del internet se vuelve indispensable para poder contar con estrategias más certeras. Dicho de otro modo, tener aproximaciones objetivas respecto al entendimiento de las preguntas propuestas, permiten definir líneas de políticas más precisas sobre las intervenciones que podría hacer el Estado a fin de lograr que las poblaciones utilicen efectivamente el servicio de internet y logren apropiarse de los beneficios derivados de ello.

Una pregunta derivada del análisis que tiene relevancia de política, y que puede derivar en estudios posteriores, es ¿debemos seguir priorizando políticas de oferta o debemos complementarlas con políticas de demanda? Si es así, ¿en dónde focalizarlas?

Por tanto, el objetivo general de la presente investigación es su contribución con la determinación de manera empírica de las causas que inciden en la brecha de uso del

internet, identificando las causas observables y aquellas no observables, aplicando para ello la descomposición de Oaxaca-Blinder.

De manera complementaria, la investigación contribuye con una revisión de la literatura sobre estimaciones de la brecha de uso del internet; así como con el planteamiento de un modelo (económico) que explica la brecha del uso del internet, y en ese sentido, proporciona un punto de partida (una base de datos inicial) con información estadística que permita ahondar en el análisis correspondiente en la medida que se logre ampliar dicha información tanto a nivel temporal como en cuenta a los niveles de desagregación que se considere.

1.3. HIPÓTESIS

Las hipótesis que se plantean buscan responder la pregunta ¿cuáles son los principales determinantes de la brecha de uso del internet urbano-rural en el Perú?. Por tanto, las hipótesis que se plantean son las siguientes:

1) HIPÓTESIS 1: Brecha de Retornos Educativos en Zonas Rurales

El sistema educativo actual contribuye a la persistencia de la brecha digital al generar retornos desiguales en habilidades digitales entre zonas urbanas y rurales. Esto sugiere la necesidad de reformas curriculares con enfoque en competencias digitales aplicadas y programas de capacitación docente adaptados al contexto rural.

2) HIPÓTESIS 2: Las Barreras Perceptivas como Amplificadores de la Brecha

Las barreras percibidas por el lado de la demanda —como el costo o la utilidad percibida del Internet— no son relevantes mientras no exista acceso, pero se convierten en factores críticos una vez desplegada la infraestructura, actuando como amplificadores del rezago digital en contextos rurales.

3) HIPÓTESIS 3: La Brecha Digital como Fenómeno Multidimensional

La desigualdad en el uso de Internet entre zonas urbanas y rurales es el resultado de un entramado multifactorial, donde:

(a) la infraestructura establece el piso mínimo de acceso,

(b) la capacidad económica define la posibilidad de adopción y factores como la educación.



CAPÍTULO 2: MARCO CONCEPTUAL Y CONTEXTO

2.1. REVISIÓN DE LA LITERATURA

Con el fin de contextualizar el tema de análisis de la presente tesis dentro de la literatura existente, es bien sabido que existe una extensa literatura que ha analizado el problema de las brechas de acceso y uso en telecomunicaciones, ya sea de manera separada como de manera conjunta, y muchas veces se les pone el rótulo general de brecha digital.

Tal como señala Park et al. (2015), estos estudios se han centrado en dos áreas; en primer lugar, la medición y cuantificación del grado en el cual ocurre la brecha digital, para lo cual se han empleado diversas técnicas estadísticas, desde indicadores sencillos (individuales) hasta indicadores compuestos y producto de su revisión de la literatura, los autores recopilan las variables comúnmente utilizadas para la medición y cuantificación, las cuales serían número de usuarios de internet, servidores de internet, teléfonos móviles, computadores y líneas telefónicas, como indicadores de infraestructura¹¹.

La segunda área de estudio señalada por Park et al. (2015), tiene que ver con la identificación de los principales determinantes asociados con la brecha digital para diferentes niveles de agregación¹². En ese sentido, los autores señalan que si bien a nivel individual las variables más utilizadas son la clase socioeconómica, género, edad y región¹³; estudios a nivel de países deben considerar otras variables explicativas,

¹¹ Los autores mencionan en este caso a Corrocher y Ordanini (2002); Vicente y Lopez (2006); Beilock y Dimitrova (2003); Kiiski y Pohjola (2002); Oyelaran y Lal (2005); Bagchi (2005); Cilan et al. (2009); Vicente y Lopez (2006); (Chinn and Fairlie, 2007; Ahn and Lee (1999); Dasgupta et al. (2001).

¹² En este caso, la literatura referida por Park et al. (2015) corresponde a Andanova (2006); Chinn y Fairlie (2007); Kiiski and Pohjola (2002) y Pohjola (2003).

¹³ Los autores señalan que ello es así, toda vez que a nivel individual las TIC no es aspecto neutral desde las perspectivas socioeconómicas.

dependiendo precisamente de las características de los países, la tecnología e incluso los enfoques de las investigaciones¹⁴.

Estudios más recientes, tales como Zhou et al. (2024), muestran un resumen de la literatura bastante exhaustivo respecto al tratamiento de la brecha digital, ya sea desde la perspectiva de la connotación que presenta la brecha digital¹⁵ (y que se refiere a las diferencias en la accesibilidad, en las habilidades digitales, y en la escala de los beneficios que se obtienen por contar con habilidades digitales); desde la perspectiva de los aspectos de medición y construcción de índice; y desde la perspectiva de los determinantes de la brecha (diferenciando diferencias en el capital físico, capital humano y capital social)¹⁶.

Por su parte, Doyar et al. (2023), presenta un resumen de la literatura referida a los determinantes del uso de las TIC, y precisan que las variables usualmente utilizadas han sido el género, el estado civil, la edad, la presencia de hijos en el hogar, el nivel de educación, el idioma, el estado laboral, los ingresos, el área de residencia, las habilidades digitales, y la posesión de equipamiento tecnológico. Un valor adicional de este trabajo se refleja en su Tabla 1, donde se esquematiza el referido resumen y se listan las variables mencionadas, mostrando la influencia de cada una de ellas en el uso de internet (los signos estimados obtenidos en cada uno de los paper revisados). Y es a partir de dicha tabla que los autores plantean los impactos esperados de las variables relevantes

¹⁴ En este extremo, los autores refieren a Crenshaw and Robison (2006); Zhao et al. (2007); Chinn and Fairlie (2007); World Bank (2006), y Billon et al. (2009).

¹⁵ Que, según los autores, se encuentra influenciado por el continuo desarrollo de las TIC y sus aplicaciones.

¹⁶ Dado que el estudio se realiza a nivel nacional, regional y provincial de China, los autores identifican indicadores asociados a brechas, para cada uno de los niveles analizados. En la Tabla 1 del paper (pg 8) se muestran dichos indicadores.

en la probabilidad de usar el internet y en la frecuencia del uso del internet. En el Anexo N° 2 del presente trabajo se muestra dichos efectos esperados.

Tsetsi y Rains (2017), a partir del análisis de las actividades (sociales o informativas) que pueden desarrollar los diversos grupos demográficos en Estados Unidos, desde los *smartphones*, examina cómo las diversas características demográficas impactan en el uso del Internet, encontrando brechas de uso no solo en lo referente al consumo de información sino también en cuanto al uso del Internet para actividades sociales. Las variables que encuentra significativas son la raza, el sexo, la edad, los ingresos y la educación. Además encuentra que los grupos minoritarios (de bajos ingresos, jóvenes, con menor nivel educativo) usan generalmente el Internet para actividades sociales; mientras que los grupos favorecidos (hombres blancos, usuarios nativos de Internet, de mayores ingresos y con mayor nivel educativo) usan el Internet para actividades de información.

Martínez-Domínguez y Mora-Rivera (2020) identifican variables socioeconómicas y demográficas que favorecen la adopción y el uso del Internet en la población rural mexicana. Prácticamente coinciden con Doyar et al. (2023) respecto a las variables que inciden en (la probabilidad) de uso del Internet.

Para el caso peruano, algunos autores han analizado la brecha digital desde la perspectiva del uso. Así, De Los Ríos (2010), a partir de la ENAHO 2007-2009, estima la brecha de uso de Internet para hogares peruanos, y encuentra que las variables más significativas que explican el uso y la brecha de uso son el ingreso del hogar, la edad, educación y sexo del jefe de hogar; la ubicación geográfica (urbana, rural, capital de distrito) y el número de perceptores de ingresos en el hogar.

Por su parte, García et al. (2011) analizan la brecha de uso del Internet sujeto a la tenencia de computadoras en Perú, utilizando datos de la ENAHO para el período 2001-2006. Los autores encuentran que las variables significativas resultan ser el ingreso del hogar, el nivel de educación del jefe de hogar, el acceso a la telefonía fija-móvil-Tv paga, el área de residencia, familiares con estudios, condición laboral del jefe del hogar y la tasa de interés promedio (como proxy del costo de acceder a créditos).

Barrantes et al. (2019) analizan la brecha digital de género en el uso de las TIC para 5 países latinoamericanos entre los cuales se encuentra el Perú. Para ello emplean la metodología de emparejamiento (*matching*) entre hombres y mujeres con características observables similares, lo cual permite descomponer la brecha de género total en un componente explicado por diferencias observables (como educación, ocupación, entre otros) y en un componente no explicado por dichas diferencias observables (y por tanto, es atribuido a factores no observables como las normas sociales o estereotipos). En este caso, las principales variables observadas que explican la brecha de uso son la edad, el nivel educativo, la presencia de menores de edad en el hogar, la ubicación rural, la lengua originaria, el nivel socioeconómico y ocupación del jefe de hogar).

Gamboa y García (2024) analizan para el caso peruano “los patrones que caracterizan las disparidades digitales que enfrentan los distintos grupos demográficos, centrándose en el uso del Internet y de los dispositivos digitales, así como la dinámica de su uso entre el 2013 y el 2022. Respecto a las dinámicas en el uso del Internet, encuentran que “las actividades más comunes son la obtención de información, comunicación y entretenimiento, aunque la obtención de información ha disminuido, mientras que el entretenimiento ha aumentado”. Y con relación al limitado uso del Internet, las variables que explican este hecho son “no tener educación formal, estar en pobreza extrema, residir en zonas rurales (para menores de 17 años), ser adulto mayor, y no tener una ocupación

vinculada al uso de Internet o dispositivos digitales”. La Tabla 2 presenta un resumen de la literatura revisada.

Tabla 2

Resumen de la literatura

Autor	Principales resultados
Tsetsi y Rains (2017).	<p>Se analiza la brecha de uso entre grupos demográficos en cuanto a tipos de actividades realizadas desde smartphones: actividades sociales (como redes sociales y Twitter) y actividades de información/noticias (como buscar noticias, salud, gobierno). También se analiza cómo el nivel de acceso a dispositivos (solo smartphone vs. acceso multimodal) modula esa brecha de uso.</p> <p>Las variables demográficas consideradas como explicativas son: raza (blancos vs. minorías), sexo (hombre vs. mujer), edad, ingreso y educación.</p>
Doyar et al. (2023)	<p>Analiza la brecha de uso del Internet, considerando tanto la probabilidad de ser usuario como la frecuencia de uso (mensual, semanal, diario). Se enfoca en desigualdades en el uso del Internet por factores como género, edad, educación, conocimiento del inglés, nivel de ingresos, localización, ocupación, etc.</p> <p>Las principales variables que explican el uso de internet o la brecha de uso, son sexo, edad, estado civil, hijos en el hogar, educación, conocimiento del inglés, ocupación (empleado, autoempleado, otros, desempleado), ingreso del hogar, ubicación geográfica (rural, urbano, capital), habilidades con PC, dispositivos con conexión a Internet.</p>
De Los Ríos (2010)	<p>Analiza la brecha de uso de internet en hogares peruanos, especialmente considerando si el hogar tiene al menos un miembro que usa internet o no. También se analiza la brecha de intensidad de uso entre trabajadores.</p> <p>Principales variables que explican el uso de internet o la brecha de uso, son: disponibilidad de cabinas públicas de internet (en lugar de tenencia de computadora), ingreso del hogar, edad, educación y sexo del jefe del hogar, número de perceptores de ingreso, frecuencia de uso de internet, y ubicación geográfica (urbano, rural, capitales de distrito).</p>
Barrantes et al. (2019)	<p>Analiza la brecha de género en el uso de las TIC en cinco países de América Latina.</p> <p>Las principales variables observadas que explican el uso de internet o la brecha de uso, son: edad, nivel educativo, presencia de menores en el hogar, ubicación rural, lengua originaria, nivel socioeconómico y ocupación.</p>
García et al. (2011)	<p>Analiza la brecha en el uso del servicio de Internet en el hogar, condicionada a la tenencia de computadora.</p> <p>Las principales variables que explican el uso de internet o la brecha de uso, son: nivel de ingreso del hogar, educación del jefe del hogar, acceso a telefonía fija, móvil y televisión pagada, tasa de interés promedio, zona geográfica de residencia, presencia de menores o personas con estudios en el hogar, condición laboral y categoría ocupacional del jefe del hogar.</p>
Martínez-Domínguez y Mora-Rivera (2020)	<p>Los hogares con mayor capital humano y físico tienen mayor incentivo a usar el internet (es decir, riqueza y educación son claves para la penetración del internet).</p> <p>La decisión de usar o no el internet depende de los beneficios esperados por estar conectados y de los costos tanto monetarios como percibidos: a mejor condición económica (acceso material y capacidad de pago), mayor uso de internet; la probabilidad de uso es mayor en mujeres, gente jóvenes y en personas con alto grado de educación; a mayores habilidades digitales, mayor uso de internet; y el mayor uso del internet se da en la región norte del país (zonas más industrializada, más desarrollada).</p>
Gamboa y García (2024)	<p>Busca identificar los patrones que caracterizan las disparidades digitales que enfrentan los distintos grupos demográficos, analizando el uso de Internet y dispositivos digitales, y su dinámica de uso entre el 2013 y 2022.</p> <p>Los principales factores limitantes del uso de internet, son: no tener educación formal, pobreza extrema, residir en zonas rurales (para menores de 17 años), ser adulto mayor, y no tener una ocupación vinculada al uso de Internet o dispositivos digitales.</p>
Park et al. (2015)	<p>Las variables que afectan o determinan la brecha digital de uso, son distribución geográfica de la infraestructura TIC, residencia en áreas remotas o con alta densidad poblacional, propiedad de dispositivos o servicios digitales (computadoras, internet o teléfonos móviles), habilidades digitales, disponibilidad de información y redes de comunicación.</p>

Elaboración propia.

2.2. LA BRECHA DE USO DEL SERVICIO DE INTERNET: CARACTERIZACIÓN

En la presente sección se describe la brecha de uso del servicio de Internet en Perú, buscando brindar una caracterización de la prestación del referido servicio, ello a partir de los resultados de la ERESTEL para diversos años -siendo la última disponible la correspondiente al año 2023-, así como de la ENAHO.

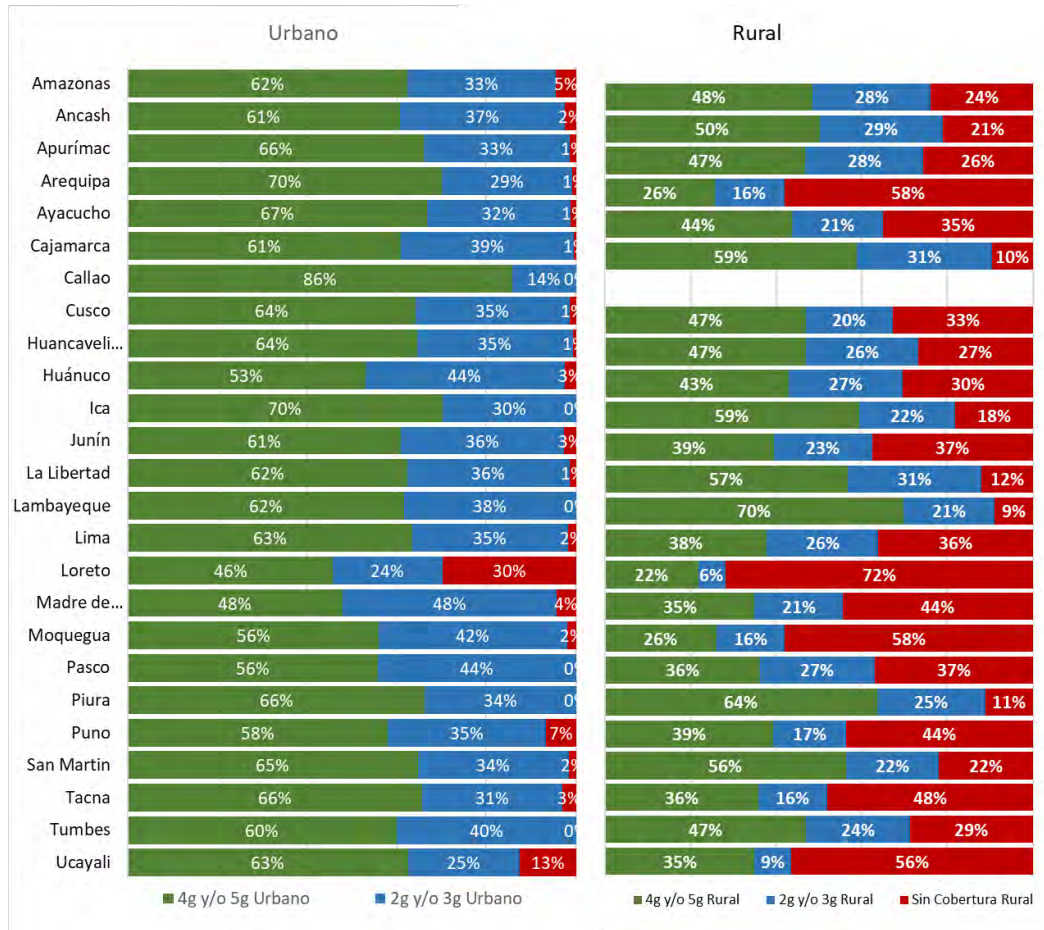
Cobertura

Si se analiza el tema a nivel regional, se puede apreciar que en su gran mayoría el servicio de internet móvil tiene mayor presencia o cobertura en las áreas urbanas de las regiones. Si bien hasta hace pocos años atrás en las zonas rurales predominaban las tecnologías 2G-3G, al cierre de 2024 la situación es distinta, pues la tecnología 4G es ahora la tecnología de acceso móvil más importante en zonas rurales.

El servicio de internet móvil (al menos tecnología 4G) tiene presencia/cobertura casi total (100%) cuando las localidades cuentan con mayor población (más de 5000 habitantes).

Figura 13

Cobertura del servicio de internet móvil (4G) a nivel regional: por segmento urbano-rural y tecnología, 2024



Elaboración propia a partir de estadísticas OSIPTEL (NRIP al 2024)

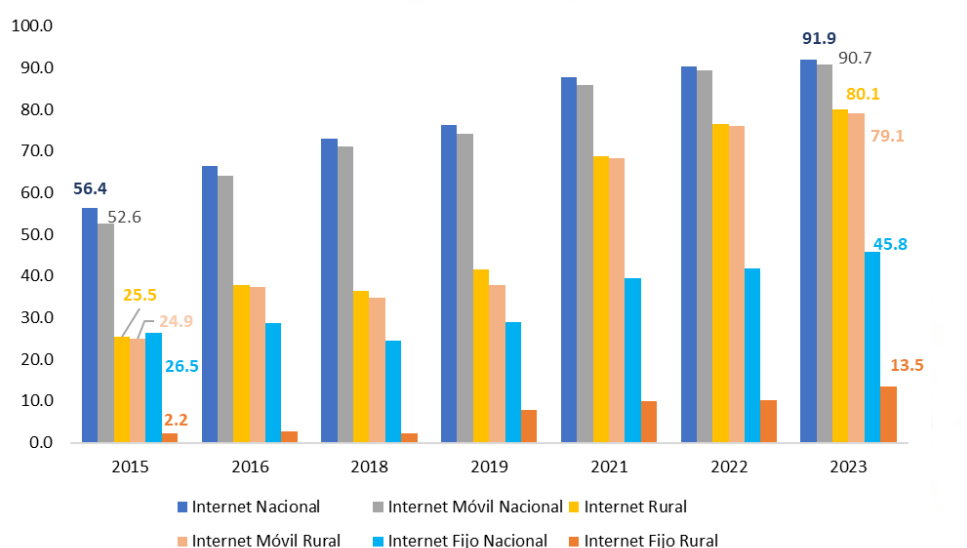
Acceso de hogares a internet

De acuerdo a la Figura 14, desde el año 2015 se ha evidenciado un incremento importante del acceso a internet en el Perú (ya sea fijo o móvil), pasando de un 56.4% en 2014 a un 91.9% en 2023. Si se desagrega esa información por ámbito geográfico, se observa que el crecimiento del acceso a internet fijo y/ o móvil en áreas rurales ha sido sustantivamente mayor. En los últimos 8 años, el acceso al servicio de Internet Móvil en áreas rurales

creció en 218% comparado con el crecimiento del Internet Móvil nacional de 73%. Más llamativo resulta el crecimiento en el acceso al Internet fijo en áreas rurales entre el 2015 y el 2023: lo hizo en 523% (aunque claro, el punto de partida fue muy bajo cercano al 2%).

Figura 14

Cobertura del servicio de Internet por segmento urbano-rural, 2015-2023

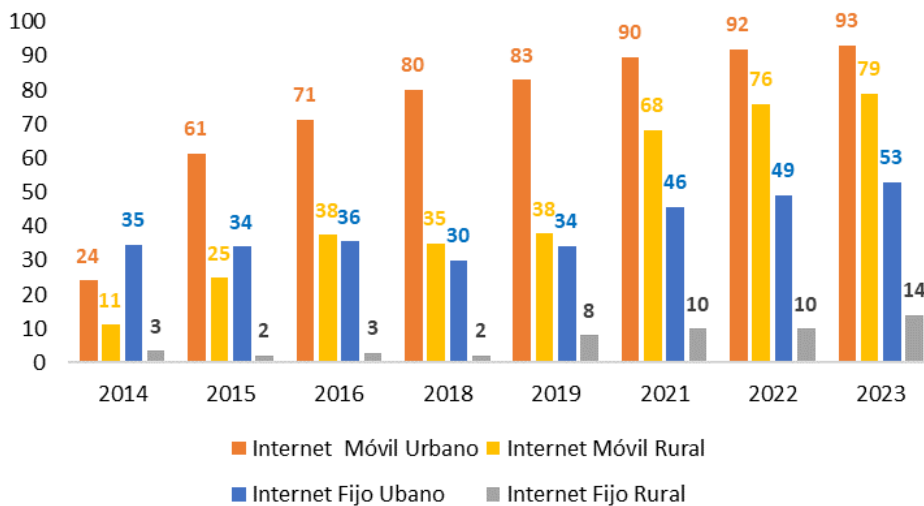


Elaboración propia a partir de estadísticas OSIPTEL (NRIP al 2024)

Desagregando por tipo de conexión (por tipo de redes de acceso) ya sea móvil o fija, se tiene que el acceso de los hogares al internet es predominantemente a través del servicio móvil. Al cierre de 2023, el 93% de los hogares tiene acceso a internet móvil en áreas urbanas y el 79% en áreas rurales, mientras que el 53% de los hogares en áreas urbanas accede al servicio de internet fijo y sólo el 14% de hogares en áreas rurales accede al referido servicio. Al respecto, la decisión sobre el tipo de acceso que preferirán los hogares, va depender del costo del servicio (costo de acceso y uso), de los atributos que ofrezca cada servicio (que claramente son distintos entre sí) y como no, de las necesidades específicas de cada familia.

Figura 15

Acceso a internet según tipo de conexión y según ámbito geográfico, 2014-2023

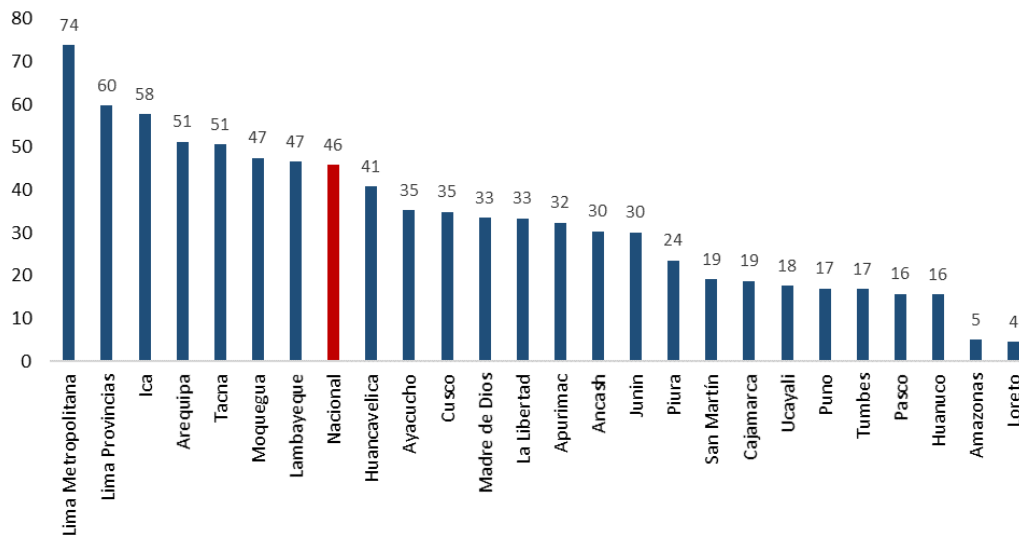


De otro lado, de la Figura 15 se observa que existe una clara brecha de acceso entre el internet móvil y el internet fijo, pero la brecha se hace más significativa cuando la comparación se hace entre áreas rurales y urbanas, y ello es mucho más notorio en el caso de internet fijo. El menor costo de despliegue de redes móviles respecto de las redes de internet fijo ha contribuido a que la brecha de acceso urbano-rural sea menor e incluso se haya ido reduciendo en el caso del internet móvil, en particular a partir del año 2021.

También resulta interesante conocer el grado de acceso al servicio de Internet fijo a nivel de regiones. De acuerdo a la ERESTEL 2023, la tasa de acceso a nivel nacional es de 46%, habiendo 6 regiones que cuentan con mayores tasas de acceso a internet fijo, destacando la región Lima (Lima Metropolitana tiene una tasa de acceso de 74%). Las regiones rezagadas en el acceso a internet son Amazonas y Loreto con tasas de 5% y 4%, respectivamente. La Figura 16 muestra el detalle para las 24 regiones del país.

Figura 16

Acceso a internet fijo por regiones, 2023

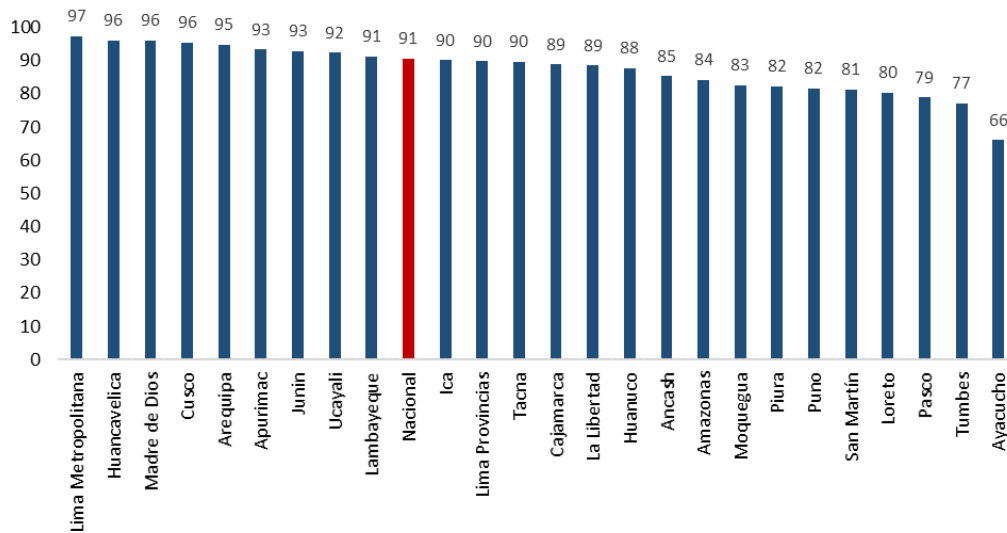


Elaboración propia a partir de la Erestel 2023 (OSIPTEL)

Por el lado del acceso al internet móvil, la Figura 17 muestra que la referida tasa de acceso a nivel nacional es de 91%, donde 9 regiones presentan tasas de acceso mayores al promedio nacional. Lima Metropolitana cuenta una tasa de acceso de 97%, seguida por Huancavelica, Madre de Dios y Cusco, todas ellas con 96%. La región con menor tasa de acceso a Internet móvil es Ayacucho con un 66%, con 11 puntos porcentuales por debajo de Tumbes.

Figura 17

Acceso a internet móvil por regiones, 2023



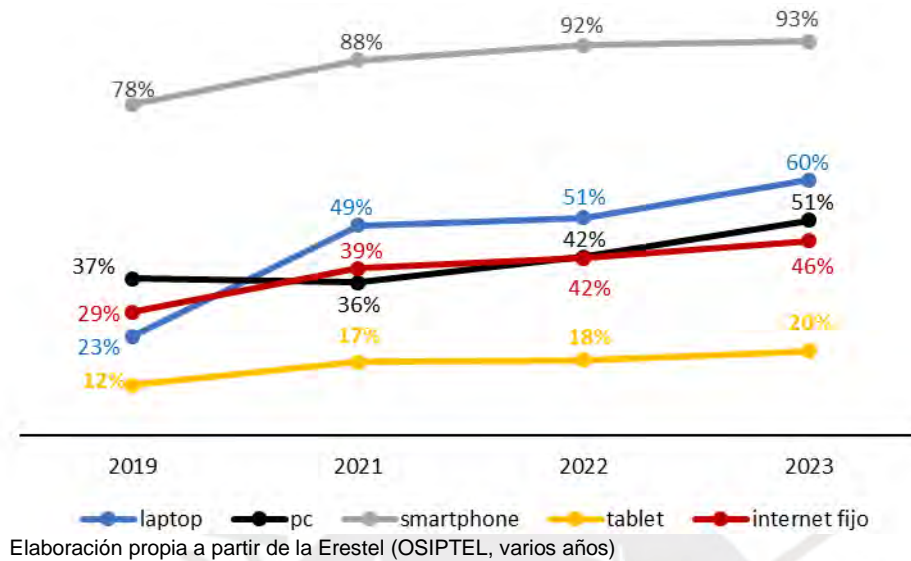
Elaboración propia a partir de la Erestel 2023 (OSIPTTEL)

Equipamiento para el acceso a internet

De otro lado, está claro que un elemento importante para la decisión de acceder al servicio de internet (fijo o móvil) se relaciona con la posesión o adquisición previa de los dispositivos que permitan hacer uso del servicio de internet. Al respecto, la Figura 18, muestra que en los últimos 3 años, el acceso a laptops, computadoras de escritorio ha mantenido una tendencia creciente y en el mismo sentido que el acceso a Internet. Entre el 2019 y el 2023, el acceso a laptop por parte de los hogares se ha más que duplicado, mientras que el acceso a computadoras de escritorio ha tenido un incremento del 37%. Al cierre del 2023 el equipo TIC más importante que permite el uso de Internet fue el Smartphone (92,8% de los hogares lo tienen) y en segundo lugar se encuentra la laptop, que desde el 2021 superó a la PC, y actualmente su tasa de acceso es de 60%.

Figura 18

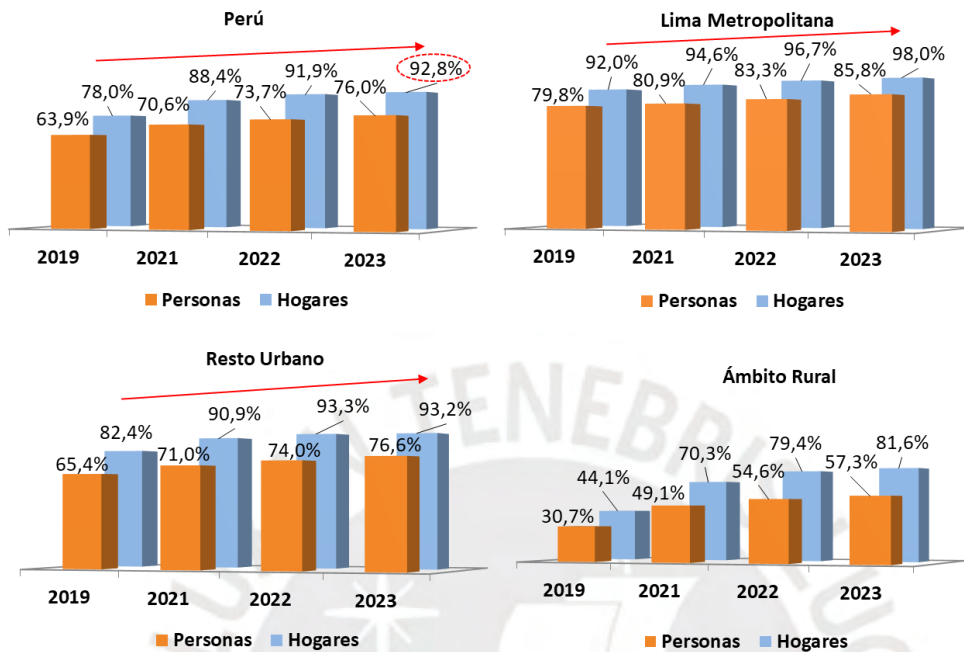
Equipamiento TIC del hogar, para uso de Internet: 2019-2023



Si se desagrega la información a nivel de ámbito geográfico, se observa que el crecimiento del acceso a Smartphone fue mayor y más notorio en el ámbito rural, pues entre 2019 y 2023 se incrementó en 85%, pasando de una tasa de 44,1% en 2019 a un 81,6% en 2023.

Figura 19

Acceso a smartphones según ámbito geográfico, 2019-2023



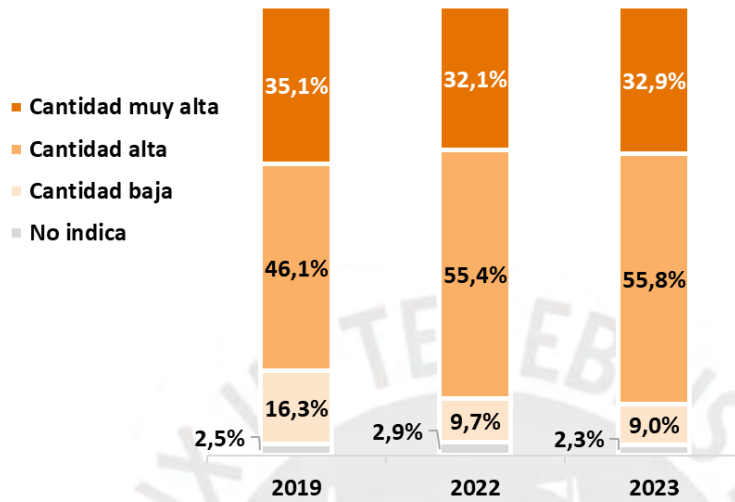
Elaboración propia a partir de la Erestel (OSIPTEL, varios años)

Percepción de atributos contratados en internet móvil

Las mejoras en atributos de los planes de internet móvil (oferta comercial) ha tenido un correlato en la percepción de los usuarios de planes móviles respecto de la mayor cantidad de datos disponibles y respecto de mayores velocidades experimentadas a la hora de usar el internet. La percepción favorable sobre mayor cantidad de datos en los planes tarifarios pasó de 81,2% en 2019 a 88,7% en 2023; y la percepción positiva sobre mejores velocidades ofrecidas en los planes de internet móvil pasó de 85,3% en 2019 a 89% en 2023.

Figura 20

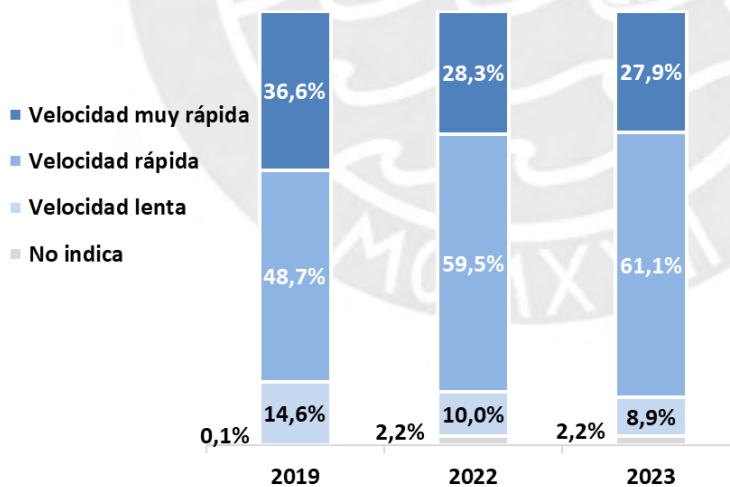
Percepción sobre la cantidad de datos del servicio de internet móvil desde el celular, 2019-2023



Elaboración propia a partir de la Erestel (OSIPTEL, varios años)

Figura 21

Percepción sobre la velocidad del servicio de internet móvil desde el celular, 2019-2023



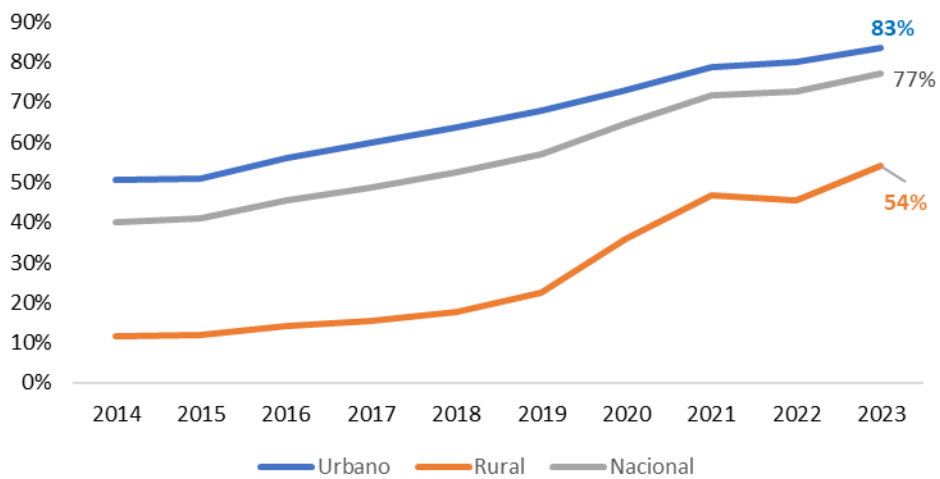
Elaboración propia a partir de la Erestel (OSIPTEL, varios años)

Uso según ámbito geográfico, 2014-2023

Si bien se evidencia al cierre del 2023 un crecimiento sostenido en el uso del internet, el crecimiento ha sido más notorio en el ámbito rural, aunque se aprecia claramente una brecha de casi 30% respecto al ámbito urbano.

Figura 22

Usos de internet según ámbito geográfico, 2014-2023



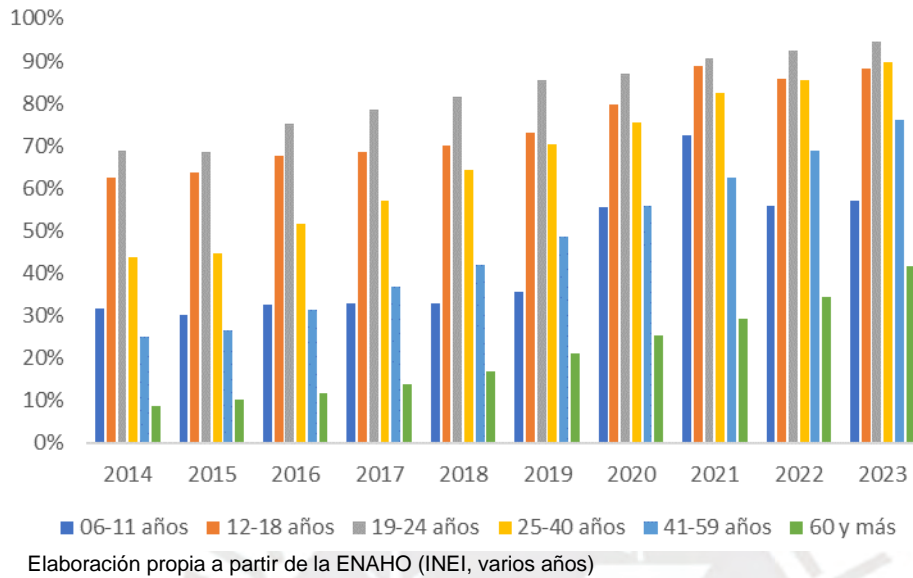
Elaboración propia a partir de la ENAHO (INEI, varios años)

Uso según edad, 2014-2023

Si bien los grupos etarios que hacen mayor uso del internet (tasas de uso entre 88% y 95%) comprenden el rango que va desde los 12 hasta los 40 años, el mayor impacto en el uso de internet en los últimos 5 años se ha dado en el grupo de personas mayores a 60 años, pues su tasa de uso casi se duplicó (pasó de 21% en 2019 a 42% en 2023).

Figura 23

Uso de internet según edad, 2019-2023

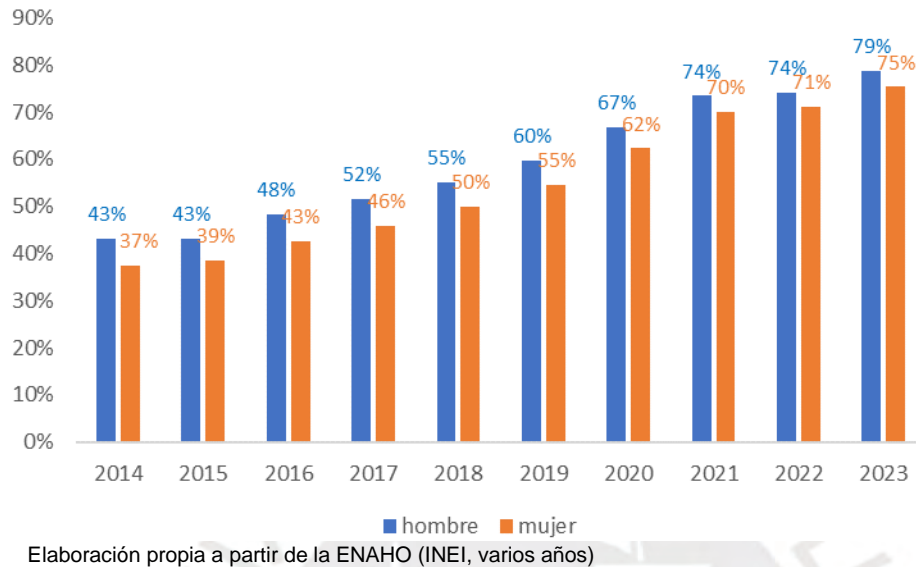


Uso por sexo

A nivel agregado el uso de internet se ha ido incrementando sostenidamente en el tiempo, pues el 2019 la tasa de uso fue de 57,1% y el 2023 fue de 77,1%. Si bien los datos muestran una brecha persistente en el uso de internet entre hombres y mujeres de alrededor 5% (al cierre del 2023 la tasa de uso de los hombres fue de 79% y de las mujeres fue de 75%), en los últimos 5 años el incremento en el uso del internet ha sido mayor en mujeres (38,3%) que en hombres (31,8%)

Figura 24

Uso del internet por sexo, 2014-2023

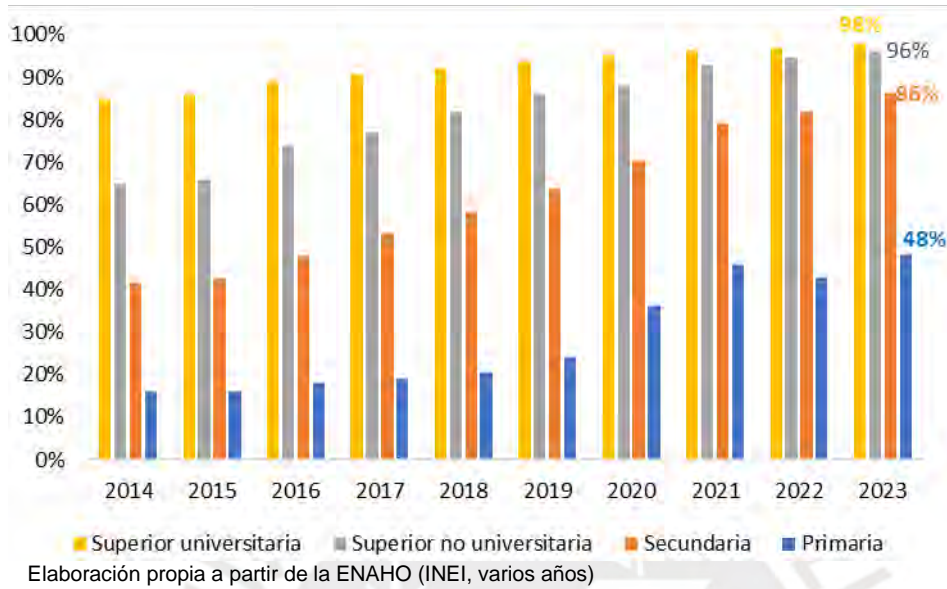


Uso por nivel educativo

Por el lado del nivel educativo de la población, al cierre del 2023, el 98% de las personas que cuentan con educación superior universitaria hicieron uso del internet. Por su parte, de la población con menos nivel de instrucción (aquellas que culminaron sus estudios de primaria) solo el 48% hizo uso del internet.

Figura 25

Uso del internet por nivel educativo alcanzado como % de la población, 2014 - 2023

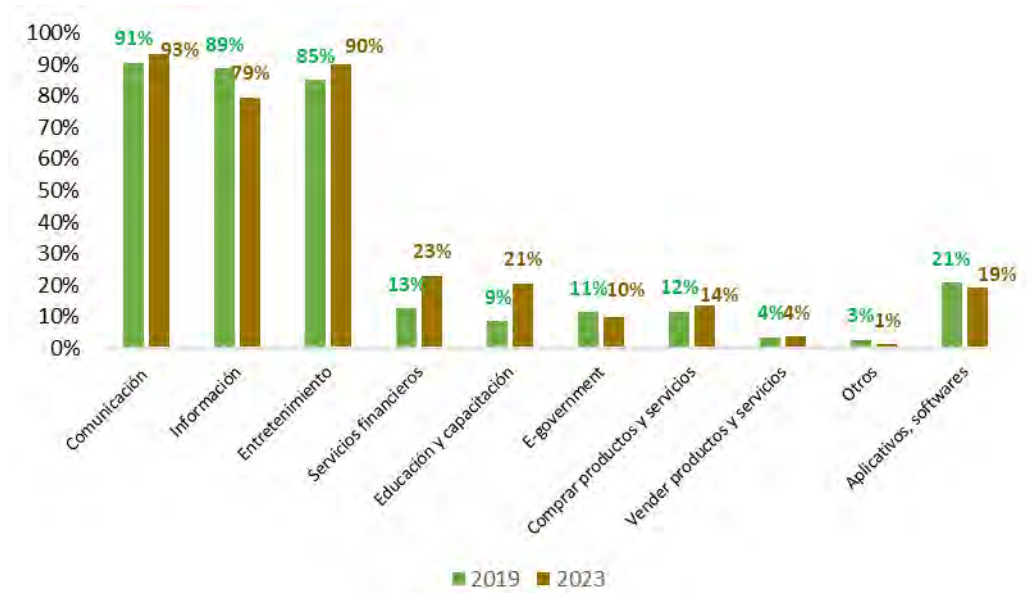


Uso por materia, interés

Los principales usos del internet son con fines de comunicación, para obtener información y para el entretenimiento. Es notable el incremento en el uso del internet para fines educativos y de capacitación en los últimos 5 años, pues se pasó de 9% en 2019 a 21% en 2023.

Figura 26

Formas de Uso del internet según materia de interés, 2014-2023



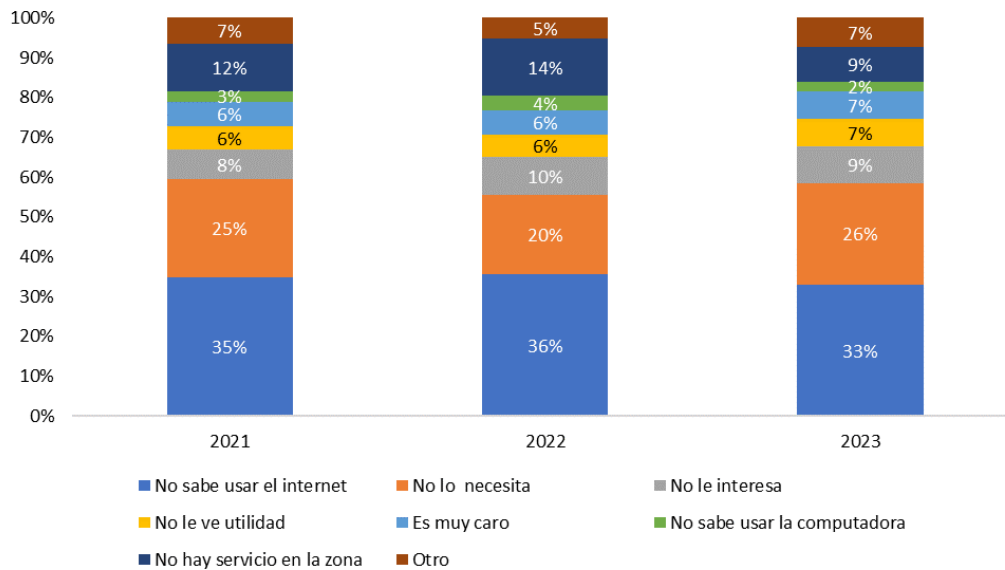
Elaboración propia a partir de la ENAHO (INEI, varios años)

Razones para no contratar internet

Un aspecto relevante de conocer, que está vinculado directamente con el objetivo de la tesis, tiene que ver con las razones detrás de la decisión de no contratar el servicio de acceso a internet, ya sea fijo o móvil. Si bien uno podría pensar que la variable precio es la principal razón (ocupa el 5to lugar de importancia), resulta bastante informativo para los fines de cualquier determinación de política pública, que las principales razones se relacionan con el desconocimiento sobre el uso del internet, la declaración de no necesitar del servicio (lo cual evidencia una falta de conocimiento sobre los beneficios esperados de acceder y apropiarse de los contenidos del internet).

Figura 27

Razones para no usar el Internet, 2021-2023



Elaboración propia a partir de la Erestel (OSIPTEL, varios años)

Siendo que el objetivo de la tesis es identificar factores que explican las diferencias en el uso del internet entre población urbana y población rural, los indicadores hasta aquí presentados brindan ciertas luces respecto a posibles causas que podrían explicar dicha brecha. Algunos resultados son consistentes con lo que la literatura ha mostrado en las secciones previas: variables como educación, edad y aspectos motivacionales son elementos importantes que inciden en el uso y tasa de uso del Internet. La identificación de grupos específicos asociados a cada variable (población menos educada, población con mayor edad) cruzado con los aspectos motivacionales detrás de su decisión de usar o no el Internet, debe permitir definir mejor las políticas públicas orientadas a, precisamente, lograr un mayor y mejor (más productivo) uso del Internet.

2.3. DETERMINANTES DE LA BRECHA DE USO DE INTERNET. EL MODELO

En la sección 2.1 se presentó un resumen de la literatura sobre el acceso y brecha de acceso digital desde la perspectiva del uso, y sobre dicha base se puede plantear una lista de variables o factores que inciden en la demanda o uso del servicio de Internet.

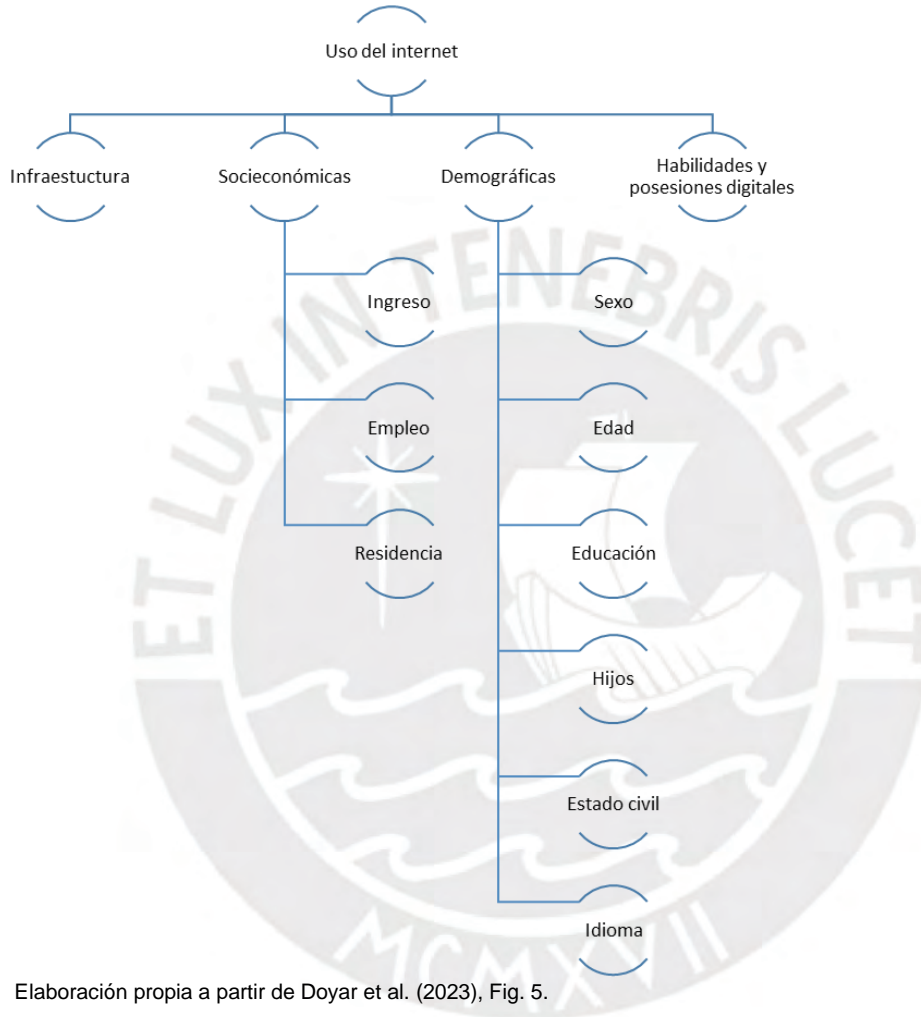
Así, la referida lista de variables socioeconómicas y demográficas que favorecen la adopción y el uso del Internet es la siguiente:

- i. Número de usuarios de internet (acceso a Internet)
- ii. Número de teléfonos móviles (acceso al servicio móvil)
- iii. Número de computadoras
- iv. Clase o nivel socioeconómico (del hogar)
- v. Sexo (del jefe de hogar)
- vi. Edad
- vii. Nivel de educación logrado (por el jefe de hogar y otros miembros del hogar)
- viii. Región de procedencia (o área de residencia)
- ix. Estado civil (del jefe de hogar)
- x. Presencia de hijos (o menores de edad) en el hogar
- xi. Idioma (o lengua materna)
- xii. Raza
- xiii. Situación laboral (del jefe de hogar)
- xiv. Ingreso del hogar
- xv. Número de perceptores de ingresos en el hogar
- xvi. Posesión de equipamiento tecnológico
- xvii. Habilidades digitales

La Figura 28 resume de manera esquemática lo listado anteriormente.

Figura 28

Variables que inciden en el Uso de Internet



Por lo tanto, el modelo base que se plantea para analizar la brecha de uso de internet urbano-rural, y más específicamente encontrar la incidencia de factores observable y no observables en la referida brecha, es el siguiente:

$$Uso\ de\ Internet = f \left(\begin{matrix} \text{variables} & \text{variables} & \text{variables} & \text{percepciones} \\ \text{económicas} & \text{socio} & \text{infraestructura} & \text{de la población} \\ & \text{demográficas} & & \end{matrix} \right) \quad (1)$$

es decir, los determinantes del uso del Internet pueden ser categorizados en cuatro grupos/componentes distintos: aspectos económicos, componente sociodemográfico, infraestructura tanto de telecomunicaciones como de servicios básicos, y un vector de percepciones de la población.

Luego se puede simplificar la notación, considerando lo siguiente:

- y , representa la tasa de uso de Internet, es una variable continua.
- X , agrupa a una serie de variables económicas (como por ejemplo, nivel de ingresos, gastos en Internet, capacidad económica, entre otros),
- Z , representa a un vector de variables sociodemográficas (edad, sexo, nivel educativo, condición laboral, etc)
- W , contiene las variables referidas a la infraestructura de telecomunicaciones y de servicios básicos, lo que puede incluir aspectos de calidad de la misma (cobertura del servicio, número de líneas o conexiones, indicadores de calidad del servicio, entre otros),
- R , agrupa variables asociadas a las percepciones de la población sobre la pertinencia/necesidad de usar el Internet (tasa de población que no usa Internet porque no sabe usarlo, tasa de población que no usa Internet porque no lo necesita o no le interesa, tasa de población que no usa Internet porque es muy caro, etc.).

Y de esta manera, la ecuación (1) pasa a ser:

$$y = f(X, Z, W, R) \quad (2)$$

Y más específicamente:

$$y_{it} = \alpha + \beta X_{it} + \delta Z_{it} + \gamma W_{it} + \psi R_{it} + \varepsilon_{it} \quad (3)$$

El detalle de las variables contenidas en cada aspecto considerado, es decir, la lista de variables explicativas toma como base la revisión de la literatura presentada, donde las variables explicativas pueden ser tanto continuas como ordinales/cualitativas.

Resulta importante mencionar que el planteamiento directo de una forma reducida se justifica en el hecho que no se está buscando estimar una función de demanda de uso de Internet, sino identificar el impacto de las características observadas y no observadas de las variables que explican de mejor manera las diferencias en el uso del Internet entre áreas urbanas y rurales.

Entonces, la teoría económica y los resultados obtenidos en diversos estudios empíricos permiten plantear los signos esperados de cada variable sobre la variable a explicar, o lo que es lo mismo, las siguientes relaciones de causalidad entre estas variables identificadas y el uso del Internet¹⁷ (en el Anexo 2 se muestran los efectos esperados de las variables relevantes que afectan el uso del Internet de acuerdo a Doyar et al. (2023)):

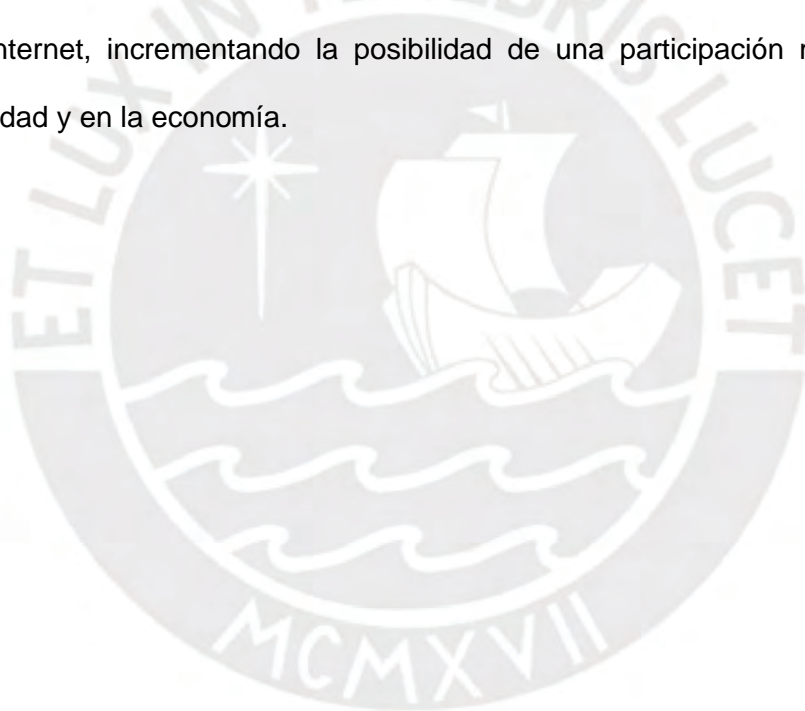
- a) Ingreso: efecto positivo, pues cuanto mayor sea el ingreso que perciben los agentes, mayor será el uso que hagan del Internet (para los agentes con menores niveles de ingreso, el efecto asequibilidad puede ser visto como una barrera).
- b) Empleo: efecto positivo, pues las personas empleadas acceden de mejor manera a las tecnologías y recursos digitales, en comparación con las personas desempleadas o subempleadas.

¹⁷ Se toma como referencia lo señalado por la Internet Society Foundation [ver <https://www.isocfoundation.org/es/2023/09/que-es-la-equidad-digital/>, búsqueda realizada el 20 de abril de 2025], que muestra los efectos relacionales entre diversas variables y lo que llaman equidad digital, siendo que esta última se entiende como el estado en el cual cada persona y comunidad cuenta con acceso a los recursos TIC y los puede aprovechar adecuadamente, para así poder participar plenamente en la sociedad, en la democracia y la economía.

- c) Región y lugar de procedencia, básicamente referido a la disparidad urbana-rural. Efecto positivo para las áreas urbanas: el mayor uso del Internet se da en áreas urbanas¹⁸.
- d) Edad: efecto negativo, pues a mayor edad, menor uso del internet (dado que en general, a la gente mayor le cuesta más adaptarse a las nuevas tecnologías e incluso en términos relativos, cuentan con menores habilidades digitales y su nivel de alfabetización digital es menor, también).
- e) Educación: efecto positivo, pues a mayor nivel educativo, mayor la tasa de uso de internet (esto tiene que ver con las mayores necesidades informacionales, con el mayor conocimiento digital y con mayores habilidades digitales).
- f) Sexo: se ha evidenciado ampliamente la brecha digital de género, con resultados a favor de los hombres: en promedio, es menos probable que una mujer utilice el Internet respecto a un hombre. De acuerdo con Alliance for Affordable Internet (2021), la probabilidad que un hombre se encuentre conectado al Internet es 21% mayor a que lo haga una mujer; y ese porcentaje se incrementa a 52% en países en desarrollo. Efecto negativo para las mujeres.
- g) Estado civil: estar casado (si existe matrimonio en el hogar) tiene un efecto positivo en la probabilidad de ser un usuario/consumidor de Internet, pero reduce la probabilidad de uso del Internet, es decir reduce la probabilidad de ser un usuario frecuente (Doyar et al, 2023).

¹⁸ La Internet Society Foundation [ver <https://www.isocfoundation.org/es/2023/09/que-es-la-equidad-digital/>, búsqueda realizada el 20 de abril de 2025] indica que para el 2022, el 82% de las personas que viven en áreas urbanas alrededor del mundo, usan el Internet, lo que equivale a 1.8 veces el porcentaje aplicable en áreas rurales.

- h) Presencia de hijos en el hogar: el efecto no es claro sobre el comportamiento de los adultos en el hogar, dado que la presencia de hijos menores en el hogar puede motivar, pero también desalentar el uso de Internet por parte de los adultos¹⁹.
- i) Lengua/idioma: efecto positivo si el idioma es el inglés (nótese que el idioma inglés es el lenguaje utilizado en el contenido del 49,2% de las páginas web a nivel mundial²⁰). Por tanto, las barreras del idioma pueden dificultar el acceso y aprovechamiento de las TIC (esto es “el acceso a los recursos digitales”).
- j) Ventajas digitales: efecto positivo dado que las personas (trabajadores, empleados, empresarios, emprendedores) con mayores ventajas digitales hacen un mayor uso del internet, incrementando la posibilidad de una participación más plena en la sociedad y en la economía.



¹⁹ Doyar et al (2023) hace referencia a que Lera-López et al. (2011) encuentran resultados inconclusos respecto al impacto de los menores en el uso de Internet en la familia.

²⁰ Dato reportado por W3Techs. (2025). De acuerdo con Doyar et al (2023), para el 21 de noviembre de 2021, el mismo W3Techs reportó que el idioma inglés era empleado como contenido en el 63,5% de las páginas web a nivel mundial.

SEGUNDA PARTE. DISEÑO METODOLÓGICO Y RESULTADOS

CAPÍTULO 3: PLANTEAMIENTO METODOLÓGICO

3.1. METODOLOGÍA

La descomposición Oaxaca-Blinder (Oaxaca, 1973 y Blinder, 1973) fue planteada originalmente para analizar la brecha salarial entre 2 grupos comparables, y así poder examinar los factores asociados con la desigualdad salarial y discriminación racial y de género en el mercado laboral.

Por tanto, este enfoque permitirá descomponer la brecha en la tasa de uso de internet entre dos grupos (por ejemplo, regiones urbanas y rurales) en dos componentes principales: uno que se atribuye a diferencias en las características observables de dichos grupos (como educación, ingreso, edad, etc.), y otro atribuido a las diferencias en características no observables directamente.

Desde su planteamiento en el año 1973, en el cual se analizaba la brecha salarial, diversos autores han empleado la referida descomposición para ampliar y complejizar el análisis inicial, así como para extender el análisis a otras áreas y poder analizar brechas y diferencias entre grupos para temáticas distintas. Ñopo (2008-i); Jann (2008); Otero (2012); Ríos Cahuas (2019) entre otros autores, presentan resúmenes sobre el uso que se le ha dado a la descomposición Oaxaca-Blinder. Así por ejemplo, una parte de la literatura ha estado orientada a expandir el análisis inicial para incorporar análisis por

quantiles²¹, variables endógenas dicotómicas (Fairlie, 1999 y Fairlie, 2006), variables endógenas censuradas²².

Como señala Jann (2008), en general la descomposición Oaxaca-Blinder puede ser empleada para estudiar las diferencias entre grupos bajo cualquier forma de la variable de resultado (ya sea continua e ilimitada), tal como por ejemplo cuando se analiza las desigualdades en la salud según el estado de la pobreza²³.

Entre algunas limitantes o críticas al enfoque Oaxaca-Blinder, Ñopo (2008-ii)²⁴ señala dos problemas potenciales de la descomposición tradicional de Oaxaca-Blinder: (i) sobreestimación del componente No Explicado (más adelante se explica este concepto) y (ii) incapacidad para abordar o analizar la distribución de las diferencias no explicadas entre los grupos evaluados.

En esa línea, Ospino et al. (2010), que hace una revisión de la literatura sobre esta metodología también señala como cuestionamientos adicionales o complementarios a los señalados por Ñopo (2008-ii), como (i) los posibles errores de especificación del modelo, (ii) la elección de variables independientes (e incluso de la forma en que ingresan al modelo), (iii) la arbitrariedad y falta de interpretación directa del componente No Explicado, (iv) posible subestimación del componente asociado a la discriminación (en el

²¹ Ñopo (2008-i) señala a Albrecht et al. (2003).

²² En este caso, Ñopo (2008-i) menciona a Bauer and Sinning (2005).

²³ Jann (2008) menciona en este caso a O'Donnell et al. (2008).

²⁴ El autor presenta estas críticas, respecto a la aplicación de la descomposición en el análisis de las brechas salariales. La primera crítica tiene que ver con un posible problema de error de especificación debido a que aun cuando existen diferencias en los resultados de las distribuciones empíricas de las características de los grupos a analizar, se termina estimando ecuaciones para ambos grupos restringiendo la comparación solo a aquellos agentes con características comparables. Por ello, se asume también que los estimadores lineales de los parámetros son válidos fuera de la muestra sobre la cual se hizo la estimación o lo que es lo mismo "fuera de los soportes de las características individuales para los cuales fueron estimados" (es decir, parámetros estimados del grupo 1 son válidos para el grupo 2). Este efecto es lo que podría llevar a sobre-estimaciones del componente No Explicado. La segunda crítica se refiere a que, debido a su construcción, la descomposición Oaxaca-Blinder es únicamente informativa acerca las diferencias promedio "no explicadas" y por tanto "no es capaz de abordar la distribución de dichas diferencias no explicadas".

mercado laboral) cuando existen diferencias marcadas entre los grupos evaluados en relación al acceso a las dotaciones que luego son recompensadas en el mercado laboral, y (v) desde la perspectiva empírica, dado que las estimaciones de los coeficientes capturan los sesgos generados a partir de problemas de información, de errores de medición de las variables y de selección de variables, resulta debatible la interpretación de estos residuales (básicamente la parte referida a la diferencia de interceptos) como aspectos asociados a una supuesta discriminación.

Para una explicación didáctica y gráfica sobre esta herramienta, puede revisarse Rahimi y Hashemi Nazari (2021).

Con relación a la aplicación de la descomposición Oaxaca-Blinder en el mercado de las telecomunicaciones, también se ha desarrollado literatura respecto a la determinación de factores observables (controlables) y no observables que explican las diferencias o la brecha digital, en cualquiera de sus diversas variantes (brecha de acceso o brecha de uso), así como las brechas donde interviene el uso TIC como variable relevante. Así por ejemplo, Vargas (2023) analiza la brecha digital de acceso en Costa Rica diferenciando el análisis por cuestiones de género y por ámbito rural-urbano. Por su parte, Liao et al. (2016) aplica la herramienta y analiza la brecha digital entre estudiantes rurales y urbanos.

Y en lo que respecta a la brecha de uso, Fairlie (2004) descompone la brecha de uso de computadoras en casa y acceso a Internet en casa, entre grupos poblacionales étnicos en Estados Unidos; y encuentra que las diferencias raciales en el uso de computadoras e Internet entre negros, mexicano-americanos y blancos se deben principalmente a diferencias en educación, ingreso y ocupación, y no a precios o diferencias escolares. Asimismo, encuentra que las barreras lingüísticas podrían influir en las bajas tasas de uso entre los mexicano-americanos.

Por su parte, Chinn y Fairlie (2006) analizan la brecha de uso en relación con la penetración de computadoras e Internet a través de un modelo de panel data entre países desarrollados y países en desarrollo (evalúan en total 161 países para el período 1999-2004). Los autores encuentran que las disparidades en ingresos, capital humano, densidad telefónica, calidad legal y desarrollo bancario explican las diferencias en la penetración de tecnología entre países desarrollados y en desarrollo.

El tema de la brecha de uso de tecnologías de la información para el teletrabajo entre trabajadores urbanos y rurales es analizado por Song et al. (2006) a través de un modelo probit y a partir de una encuesta aplicada el 2001 a la Universidad de California en Los Angeles. Los autores encuentran que el acceso a la banda ancha tiene un gran impacto en la probabilidad de teletrabajo, reduciendo la brecha entre los mercados laborales urbanos y rurales. Asimismo, a medida que el acceso a banda ancha mejora en áreas rurales, la brecha salarial entre zonas urbanas y rurales podría reducirse, permitiendo que trabajadores altamente calificados se muden a áreas más remotas.

Whitacre (2015) analiza la brecha de adopción de banda ancha entre los hogares metropolitanos y no metropolitanos (como proxys de áreas urbanas y rurales, respectivamente). Los autores estiman un modelo logit tradicional para los períodos 2003 y 2011, y sobre el referido modelo aplica una descomposición no lineal. En primer lugar, hallan que la disponibilidad de infraestructura de banda ancha en áreas no metropolitanas (léase, rurales) explica aproximadamente el 38% de la brecha de adopción de banda ancha entre zonas metropolitanas y no metropolitanas. Luego, encuentra que a pesar de los esfuerzos en infraestructura, la brecha de adopción mostró una mejoría parcial a lo largo de los años, con una disminución de la brecha en el periodo de 2003 a 2011. Sin embargo, las diferencias estructurales en características como nivel educativo y de ingresos se constituían como barreras significativas para lograr una adopción equitativa.

En ese sentido, su recomendación es que aunque la provisión de infraestructura es crucial, se debe poner más énfasis en fomentar la demanda de banda ancha en áreas rurales, por lo que programas de alfabetización digital, subsidios para la propiedad de computadoras y servicios de banda ancha se hacen indispensables.

Galperin & Arcidiacono (2021) analizan las diferencias en la participación laboral y los tipos de empleo para 5 países de América Latina (incluido el Perú), considerando para ello la brecha digital de género en 2018. Para ello emplean una regresión logit multivariable y aplican la descomposición de Yun (que es una extensión del Modelo Oaxca y Blinder para modelos no lineales). Los autores encuentran que las tasas de empleo más altas entre los hombres explican una parte significativa de la brecha digital de género en los países de América Latina. Asimismo, los resultados evidenciaron que, si bien hubo avances en igualdad de género, los estereotipos y la legislación insuficiente siguen limitando el acceso de las mujeres a sectores intensivos en TIC. Por ello, recomiendan que las políticas de inclusión digital deben integrarse con estrategias para reducir las desigualdades laborales, promoviendo la capacitación digital para mujeres de bajos ingresos.

Por lo tanto, con el planteamiento metodológico, es decir con la descomposición Oaxaca-Blinder, se busca cuantificar y explicar la brecha (las diferencias) en el uso de internet (en la apropiación del Internet) entre zonas urbanas y zonas rurales del Perú, en el período 2018-2023. Para ello el modelo a estimar es el que se representa en la ecuación (3) de la sección 2.3

Técnicamente, esto se realiza estimando una regresión lineal separada para cada grupo, donde la variable dependiente es la tasa de uso de internet, y las variables independientes son los factores que teóricamente determinan ese uso. La comparación entre ambas

regresiones permite cuantificar cuánto de la brecha se debe a diferencias en las características observables promedio entre los grupos (como por ejemplo el nivel de educación, el nivel de los ingresos, la edad, etc.) y cuánto se debe o se asocia a diferencias en cómo esas características afectan el uso de internet, es decir a diferencias en los retornos a esas características.

Lo que permite este enfoque es analizar factores estructurales (infraestructura, educación) y barreras no económicas (cognitivas, actitudinales) que persisten incluso controlando por variables tradicionales, y que afectan el uso del internet.

En el presente caso, la variable de interés es la tasa de uso de internet a nivel regional, y se estimará una regresión lineal de la forma:

$$y_i = \beta_0 + \sum_{k=1}^K \beta_k x_{ik} + \varepsilon_i$$

donde:

- Y_i : variable dependiente (por ejemplo, uso de internet en la región)
- x_{ik} : vector de características observables (educación, ingresos, pobreza, etc.)
- β_k : vector de coeficientes estimados
- ε_i : término de error

Luego, separamos por grupos:

- Grupo U: población urbana (Regiones con alta urbanidad)
- Grupo R: población rural (Regiones con alta ruralidad o baja urbanidad)

Sea \bar{Y}^U y \bar{Y}^R los promedios de uso de internet en los grupos U y R, respectivamente.

También definimos \bar{x}^U , \bar{x}^R , β^U , β^R como los promedios de características y coeficientes por grupo. La expresión básica de la descomposición, bajo el entendido que uno de los

grupos enfrenta discriminación, es la siguiente (en el Anexo 3 se muestra la derivación de descomposición Oaxaca-Blinder, hasta llegar a una expresión consistente con la forma en que el paquete econométrico Stata muestra los resultados de la referida descomposición):

$$\Delta \bar{Y} = \underbrace{(\beta_o^U - \beta_o^R)}_B + \underbrace{\sum_{j=1}^K \beta_j^R (\bar{x}_j^U - \bar{x}_j^R)}_D + \underbrace{\sum_{j=1}^K (\beta_j^U - \beta_j^R) \bar{x}_j^R}_C + \underbrace{\sum_{j=1}^K (\beta_j^U - \beta_j^R) (\bar{x}_j^U - \bar{x}_j^R)}_I$$

..... (4)

donde:

- El componente B es atribuido a las diferencias básicas entre grupos, e incluye los efectos de las variables desconocidas que no están incluidas en el modelo.
- El componente D hace referencia a los cambios esperados medios en las zonas rurales (grupo R) cuando se le comparan con (cumplen con) el nivel de las covariables de las zonas urbanas (grupo U). Es decir, este componente es la parte de la brecha que se explica por las diferencias de grupo en el nivel de las variables explicativas observables (es el componente explicado). Es conocido como el “efecto (de las) dotaciones” de cada grupo.
- El componente C es una parte de la brecha que representa un cambio esperado medio en el grupo rural cuando se le compara con (cumple con) los coeficientes de regresión del grupo urbano. Por tanto, hace referencia a la parte de la brecha que es causada por el efecto diferencial de las variables observables (este efecto diferencial se traducen en los β de cada regresión) sobre los resultados en los 2 grupos que se comparan. De esta manera esta parte de la brecha no puede ser explicada por el nivel de las variables explicativas que son observables. Es el componente no explicado; y se le conoce como el “efecto (de los) coeficientes”.

- El cuarto componente I hace referencia a la interacción debido al efecto simultáneo de las diferencias en las dotaciones y en los coeficientes.

Si ahora fuera el caso que los dos grupos enfrentan discriminación, es posible definir un vector de coeficientes no discriminatorios β_j^* , tal que ambos grupos U y R deseen alcanzarlo, donde se define β_j^* , tal que: $\beta_j^U \geq \beta_j^* \geq \beta_j^R$, o $\beta_j^U \leq \beta_j^* \leq \beta_j^R$. Entonces:

- Si $\beta_j^U > \beta_j^* > \beta_j^R$, existe discriminación positiva a favor del grupo U (y discriminación negativa en contra de grupo R);
- Si $\beta_j^U < \beta_j^* < \beta_j^R$, existe discriminación positiva a favor del grupo R (y discriminación negativa en contra de grupo U).

Trabajando la ecuación (4) se llega a

$$\Delta \bar{Y} = \underbrace{\sum_{j=1}^K \beta_j^* (\bar{x}_j^U - \bar{x}_j^R)}_{\substack{\text{Parte Explicada o} \\ \text{Efecto Dotación}}} + \underbrace{(\beta_0^U - \beta_0^R) + \sum_{j=1}^K (\beta_j^U - \beta_j^*) \bar{x}_j^U + \sum_{j=1}^K (\beta_j^* - \beta_j^R) \bar{x}_j^R}_{\substack{\text{Parte No explicada} \\ \text{(incluye el efecto Discriminación)}}} \dots\dots\dots (5)$$

donde:

- Parte Explicada (Efecto Dotación): es la parte de la brecha que es explicada por las diferencias en los niveles de las variables o características observadas.
- Parte No Explicada: es la parte de la brecha debido a diferencias en los coeficientes.

En primer lugar, se tiene la diferencia entre los interceptos de los grupos $(\beta_0^U - \beta_0^R)$, que recoge el efecto de las variables no incluidas en el modelo y que según el enfoque Oaxaca-Blinder se asocia a efectos discriminación o a restricciones estructurales. En segundo lugar, recoge las diferencias de los coeficientes de los grupos (β_j^U, β_j^R) respecto del coeficiente no discriminatorio β_j^* . Por lo tanto, captura las diferencias en

el nivel de las variables no observables, así como sus efectos diferenciales (discriminatorios). Nótese que si todas las covariables no observadas se midieran y por tanto se incluyeran en el modelo, este componente únicamente comprendería la diferencia de los coeficientes de cada grupo respecto del coeficiente no discriminatorio. Este componente determina la porción no explicada de la brecha. En el caso que se está analizando, se atribuye a diferencias en los retornos a las características observables, o a factores no observados (e.g. fallas de mercado, discriminación, eficiencia en el uso de recursos, etc).

Finalmente, bajo ciertos supuestos (ver Anexo 3) la ecuación (5) se puede reexpresar de la siguiente manera (esta manera de representar la descomposición es consistente con la nomenclatura que usa el Stata):

$$\Delta \bar{Y} = \underbrace{\sum_{j=1}^K \beta_j^U (\bar{x}_j^U - \bar{x}_j^R)}_{\text{Explicado}} + \underbrace{(\beta_0^U - \beta_0^R) + \sum_{j=1}^K (\beta_j^U - \beta_j^R) \bar{x}_j^R}_{\text{No Explicado}} \dots\dots (6)$$

donde:

- El componente Explicado: es la combinación de los componentes D e I en ecuación (4). Si bien en muchos textos a este componente se le denomina el componente explicado, de hecho una parte del mismo (parte I) contiene la diferencia simultánea de los coeficientes y de los niveles de las covariables en ambos grupos. Por eso, si se desea el componente explicado puro (y crudo), es mejor usar la descomposición bajo la forma de la ecuación (4). Está claro que la ecuación (6) es una forma específica de la ecuación (4) donde los componentes D e I han sido integradas de la siguiente manera:

$$\sum_{j=1}^K \beta_j^R (\bar{x}_j^U - \bar{x}_j^R) + \sum_{j=1}^K (\beta_j^U - \beta_j^R) (\bar{x}_j^U - \bar{x}_j^R) = \sum_{j=1}^K \beta_j^U (\bar{x}_j^U - \bar{x}_j^R)$$

- El componente No Explicado: corresponde a los componentes B y C de la ecuación (4). Surge del efecto diferencial de las variables observables, pero también del efecto diferencial (recogida en los coeficientes) y niveles de las variables no observables. Este componente determina la parte no explicada de la brecha.

Si consideramos que las regiones de alta conectividad tienen mayor educación y mayor ingreso promedio. Entonces, una parte de la diferencia en el uso de internet se puede "**explicar**" por esas diferencias observables. Pero si aún después de controlar por esas variables, sigue habiendo una diferencia considerable, entonces esa parte "no explicada" podría deberse a falta de habilidades digitales, barreras estructurales o discriminación territorial, fallas de mercado (e.g. aversión o prejuicios por el uso de internet) y a infraestructura digital no observada (percepción de calidad de red).

3.2. DESCRIPCIÓN DE VARIABLES Y DATOS

Las fuentes de datos o información son las encuestas ENAHO y ERESTEL. Asimismo, se ha recopilado diversa información sobre datos económicos publicados por el portal web del INEI (<https://www.inei.gob.pe/estadisticas-indice-tematico/>), así como de la información que las propias empresas operadoras de servicios públicos de telecomunicaciones reportan al OSIPTEL en el marco de la NRIP, como por ejemplo, datos de cobertura, conexiones, inversiones, etc.

Sobre la base de las referidas encuestas, es posible recolectar información referida a las características socioeconómicas, demográficas, así como de indicadores TIC, a nivel de regiones/departamentos.

Lamentablemente no ha sido posible incluir variables o indicadores que recojan directamente el estado o la calidad del servicio de Internet (como velocidad, latencia, etc.) debido a la falta de información disponible para todo el período analizado.

Se harán estimaciones de panel departamentales para el período 2018-2023.

Si bien es posible encontrar alguna información a nivel de individuos, hogares o localidades, e incluso se podría analizar la brecha en el uso de Internet a nivel individual (para lo cual se deberían emplear variables distintas a las consideradas en el presente estudio²⁵); la unidad de análisis principal son las regiones o departamentos del Perú (se cuenta con 25 regiones observadas por 6 años, n=150 observaciones).

Esta decisión se justifica porque el objetivo central es caracterizar las brechas departamentales en el uso de internet y entender los factores estructurales, sociales y económicos que las explican. Se busca pasar de un análisis individual a uno territorial, centrado en las diferencias persistentes entre espacios urbanos y rurales en el Perú.

Si bien las encuestas base (como la ENAHO o la ERESTEL) trabajan con hogares y personas, para este estudio se ha construido una base de datos agregada a nivel regional (por departamento) que resume variables relevantes como tasas de uso, barreras percibidas, infraestructura disponible, niveles de pobreza, informalidad y educación, entre

²⁵ Por ejemplo, la variable ingreso a nivel regional se mide con el PBI departamental, mientras que si se hiciera el análisis a nivel individual se debería considerar el ingreso familiar o ingreso individual. Asimismo, la variable infraestructura a nivel regional es una variable cuantitativa continua, mientras que a nivel individual debería ser una variable cualitativa dicotómica o *dummy* (1 si tiene y 0 si no tiene).

otros. En consecuencia, los promedios regionales son el insumo sobre el cual se realiza la descomposición Oaxaca-Blinder.

Esta agregación puede esconder heterogeneidades internas dentro de cada región. Sin embargo, esta limitación se compensa en parte con el enfoque territorial del estudio, que busca proporcionar evidencia a nivel de políticas públicas regionales y nacionales, donde muchas decisiones se toman bajo criterios agregados. Naturalmente, como recomendación, la tesis dejará abierta la posibilidad de extender el análisis hacia niveles más desagregados (provincias, distritos, centros poblados) en trabajos futuros siempre que exista disponibilidad de datos (algo que actualmente es restrictivo). Asimismo, también se puede asumir que en términos de brechas de uso la heterogeneidad de la unidad de análisis no es tan problemática. Por ejemplo, la regresión controla por conectividad, infraestructura básica e ingresos, por lo que es muy probable que estas variables tengan el mismo efecto de hacer más efectivo el uso si se analizan personas, hogares, centros poblados, distritos y/o provincias.

Además, tómesese en cuenta que las medidas de política que implementa el estado son agregadas, es decir se realizan, por lo general, a nivel nacional e incluso a nivel regional. Por tanto, el análisis y recomendaciones derivadas a nivel individual, implicaría sugerir la implementación de políticas públicas o intervenciones públicas atomizadas; las cuales además serían muy costosas en su construcción, identificación e implementación.

Tratamiento de las covariables

Uso del Análisis de Componentes Principales (PCA)

Para enriquecer el análisis y optimizar la capacidad explicativa del modelo sin comprometer los grados de libertad, las variables consideradas fueron sometidas a un proceso de reducción dimensional mediante PCA. El análisis PCA es utilizado en estudios sobre brecha digital, al respecto se puede revisar Martínez-Domínguez y Mora-Rivera (2020), Ozcan y Baycan (2023), Park et al. (2015), Bruno et al. (2023) y Corrocher (2002), entre otros.

Por tanto, esta metodología permitió sintetizar la información contenida en múltiples indicadores correlacionados en un conjunto reducido de componentes ortogonales, conservando la máxima varianza de los datos originales.

- **Infraestructura de Internet**

El análisis se centra en variables clave de conectividad digital

- *th_tm*: tasa de acceso en hogares a telefonía móvil.
- *th_int*: tasa de acceso en hogares a internet según ENAHO.
- *th_conex*: tasa de acceso en hogares a internet según ERESTEL.

La construcción del índice de conectividad a internet mediante PCA (*pca_inter*) con las variables estandarizadas de acceso a telefonía móvil (*z_th_tm*), acceso a internet (*z_th_int*) medido por la ENAHO y acceso a internet en el hogar (*z_th_conex*) medido por la ERESTEL se fundamenta en lo siguiente. En primer lugar, estas tres variables capturan dimensiones de la conectividad digital residencial: mientras *z_th_int* y *z_th_conex* miden

directamente el acceso a internet en los hogares, $z_{th_{tm}}$ actúa como proxy de la infraestructura de telecomunicaciones disponible, dado que la telefonía móvil es la puerta de entrada más común al internet, sobre todo, en contextos con baja penetración de internet fija. El KMO de 0.74 confirma que esta combinación de variables tiene una adecuación muestral aceptable para el PCA.

Si bien se cuenta con variables como th_{tf} , th_{tvcs} , pen_{lmov} y pen_{intf} , la exclusión de variables como acceso a la telefonía fija (th_{tf}) o televisión de paga (th_{tvcs}) responde a que, conceptualmente, estas métricas no están directamente vinculadas al ecosistema de internet: la telefonía fija refleja una tecnología en declive con baja penetración en zonas rurales, mientras que la televisión de paga es un servicio de entretenimiento que no necesariamente correlaciona con la capacidad de acceso a internet. Asimismo, en el mercado peruano, la expansión del acceso a internet se da a través de monoproductos, más que en la forma de empaquetamiento. Adicionalmente, se descartaron los indicadores de penetración reportados por las empresas (como líneas de internet fijo o móvil) porque agregan datos residenciales y comerciales, lo que introduciría ruido en la presente tesis que mide la tasa de uso de internet en la demanda residencial.

- Calidad

Dada la información disponible, el análisis de PCA se centró en

- $reclamos_{100h}$: reclamos por cada 100 habitantes.
- p_{ccpp_cobmov} : porcentaje de centros poblados con cobertura de internet móvil²⁶.

²⁶ Debe precisarse que mediante Resolución de Consejo Directivo N° 151.2023-CD/OSIPTTEL, el OSIPTTEL cambio su enfoque regulatorio de cobertura del servicio móvil, pasando de un esquema dicotómico a nivel de centro poblado (“tiene cobertura” o “no tiene cobertura”) a un esquema de reporte libre de manchas de (área con) cobertura. Asimismo, el nuevo enfoque define dos tipos de áreas de cobertura: Cobertura Garantizada (área donde las empresas declaran ofrecer un buen nivel de señal tal que la prestación es de buena calidad) y Capacidad Adicional de Red (área donde la prestación del servicio puede ser aceptable, pero es posible la existencia de ciertas fallas debido a aspectos geográficos o a aspectos técnicos de la propagación de la señal). A fin de guardar coherencia y consistencia en la

Estandarizando las variables, se obtuvo un KMO de 0.5, que implica que el índice se encuentra en el límite inferior de aceptabilidad. A pesar de ello, continuar usando el PCA presenta ventajas conceptuales y metodológicas frente a construir otra forma de índice como por ejemplo usar un promedio simple. En primer lugar, el PCA asigna ponderaciones óptimas a cada variable según su contribución a la varianza compartida, lo que es particularmente relevante cuando las variables tienen distintas escalas o unidades de medida subyacentes (como en este caso, donde se combina densidad de reclamos con cobertura móvil). A diferencia de un promedio simple -que asume que ambas variables contribuyen por igual-, el PCA identifica automáticamente si una variable debería tener mayor peso en el índice compuesto. Adicionalmente, aunque la correlación bivariada entre las variables sea moderada (como sugiere el KMO de 0.5), el PCA sigue siendo preferible porque captura no solo la correlación lineal directa, sino también patrones multivariados que podrían no ser evidentes en un análisis por pares.

Así pues, se generó el índice `pca_cali` que mide en escala estandarizada el nivel de calidad en la región, donde valores mayores indican mejor calidad en términos de mayor cobertura de la señal de internet móvil y menor cantidad de reclamos.

- Infraestructura básica

Para analizar el impacto de la infraestructura básica en el uso de internet, se agrupa cuatro variables clave que están relacionadas con servicios domiciliarios:

- `th_rpag`: tasa de acceso en hogares a red pública de agua.

serie de centros poblados con cobertura, para el período 2023 se ha considerado los centros poblados que cuentan con cobertura garantizada, según declaración de las propias empresas operadoras (en el anterior enfoque la supervisión de los indicadores de calidad se realizaba en las áreas que contaban con cobertura; mientras que en el nuevo enfoque la supervisión de los indicadores de calidad se realiza en las áreas con cobertura garantizada).

- th_rpal: tasa de acceso en hogares a alcantarillado.
- th_gn: tasa de acceso en hogares a gas natural.
- th_elec: tasa de acceso en hogares a electricidad.

Justificación conceptual: el acceso a estos servicios básicos es un requisito fundamental para el uso de internet.

Se estandarizaron las variables de infraestructura básica y se aplicó un PCA, obteniendo un adecuado ajuste muestral ($KMO > 0.7$) y un componente principal según el criterio de Kaiser ($\text{autovalor} > 1$) que explica una parte relevante de la varianza, generándose el índice pca_servbas que mide en escala estandarizada el nivel de dotación de infraestructura, donde valores mayores indican mejor provisión de servicios básicos.

- Edad

El análisis busca identificar patrones subyacentes en la estructura demográfica de las personas y hogares que puedan afectar el uso de internet. Así, se construyó un índice de 'Edad regional', llamado pca_edad, considerando la siguientes covariables:

- pobe_3a18: tasa de población en edad escolar (3 a 18).
- th_mees: tasa de hogares con al menos un miembro en edad escolar.
- pobam_60: tasa de población adulta mayor (60 a más).
- th_mam: tasa de hogares con al menos un miembro adulto mayor.
- av_edjh: promedio de edad del jefe/a de hogar.

El valor KMO de 0.66, si bien se encuentra en el rango medio, supera el umbral mínimo de 0.5, indicando que las variables comparten suficiente varianza común para justificar el

análisis factorial. Esto se complementa con el screen plot, donde se observa un único componente con autovalor mayor a 1 (según el criterio de Kaiser), el cual explica una proporción significativa de la varianza total (76.9%), demostrando la capacidad del PCA para sintetizar eficientemente estas dimensiones demográficas interrelacionadas en un índice unificado.

La inclusión de estas variables específicas responde a la necesidad de capturar de manera integral la estructura etaria de los hogares desde múltiples perspectivas: mientras *pobee_3a18* y *th_mees* reflejan la presencia de población joven, *pobam_60* y *th_mam* representan el extremo opuesto del espectro demográfico, y *av_edjh* actúa como un indicador sintético de la madurez del hogar. El PCA permite ponderar automáticamente estas contribuciones según su varianza compartida, evitando los sesgos arbitrarios de un promedio simple donde todas las variables tendrían igual peso.

También es importante señalar que la inclusión simultánea de variables que miden tanto la estructura etaria poblacional (*pobee_3a18*, *pobam_60a*) como su manifestación en los hogares (*th_mees*, *th_mam*) responde a una estrategia para construir un indicador más robusto y representativo de la edad a nivel departamental, ya que este enfoque permitiría:

- Reducir la heterogeneidad: al capturar no solo la disponibilidad de grupos etarios en la población general, sino también su distribución efectiva en los hogares. Por ejemplo, un departamento podría tener alta tasa de población en edad escolar (*pobee_3a18*), pero si estos jóvenes están concentrados en pocos hogares (*th_mees* baja), el impacto real en las dinámicas familiares (y por ende, en el uso de internet) sería distinto. El PCA pondera automáticamente estas disparidades
- Mitigar sesgos geográficos: En zonas rurales, la cohabitación intergeneracional es más común (adultos mayores viviendo con niños), lo que hace crucial combinar

ambas perspectivas. El índice resultante sintetiza esta complejidad en una métrica unificada que refleja mejor la realidad demográfica multidimensional.

- Aumentar la validez ecológica: Al integrar variables de hogar (*av_edjh*), se controla la posible desconexión entre datos poblacionales agregados y las unidades familiares donde realmente ocurre la adopción tecnológica. Esto es clave para estudios de brecha digital, donde el contexto inmediato del hogar determina el acceso y uso de internet.

Así, el índice no solo es estadísticamente relevante, sino también conceptualmente más completo que alternativas basadas en una sola variable. En términos interpretativos, el índice de edad (*pca_edad*) sintetiza el continuo demográfico regional de la siguiente manera: Valores altos del índice reflejan departamentos con una estructura etaria mayor, caracterizada por mayores proporciones de población adulta mayor (*pobam_60a*), mayor presencia de hogares con miembros ancianos (*th_mam*) y jefes de hogar de edad de mayor edad (*av_edjh*). Simultáneamente, estos valores altos corresponden a menores tasas de población en edad escolar (*pobee_3a18*) y menor presencia de hogares con niños/adolescentes (*th_mees*). Por el contrario, valores bajos del índice identifican departamentos con estructuras demográficas más jóvenes, donde predominan los hogares con menores en edad escolar y jefes de hogar con menor edad, junto con una reducida proporción de adultos mayores.

- Educación

Las variables consideradas fueron:

- *tp_eduap*: tasa de población que no supera la educación primaria.
- *th_medub*: tasa de hogares con algún miembro que no supera la educación primaria.

- *tp_edsup*: tasa de población que tiene algún nivel de educación superior.
- *th_medus*: tasa de hogares con algún miembro con algún nivel de educación superior.
- *aedujh*: años de educación del jefe de hogar.

La construcción del índice de Educación Regional mediante Análisis de Componentes Principales (PCA) con las cinco variables estandarizadas de capital educativo se sustenta en su capacidad para sintetizar de manera óptima las múltiples dimensiones de la educación a nivel departamental. El alto valor KMO de 0.89 confirma que estas variables comparten una varianza sustancial, lo que valida la pertinencia del análisis factorial. El scree plot, junto con la regla de Kaiser (autovalor > 1), indica que un único componente principal explica la mayor parte de la varianza (94.6%), demostrando la eficacia del PCA para reducir la dimensionalidad sin pérdida significativa de información.

La selección de variables abarca tanto indicadores poblacionales (*tp_eduap*, *tp_edsup*) como de hogares (*th_medub*, *th_medus*), complementados con los años de educación del jefe de hogar (*aedujh*). Este enfoque dual —población y hogares— permite capturar la heterogeneidad educativa de manera integral: mientras *tp_eduap* y *th_medub* reflejan carencias educativas (no superar primaria), *tp_edsup* y *th_medus* miden logros avanzados (educación superior), y *aedujh* actúa como un indicador sintético de la acumulación de capital humano en el liderazgo del hogar.

Valores altos del índice representan departamentos con mayor capital educativo (alta *tp_edsup*, alta *aedujh*), mientras que valores bajos reflejan predominio de población y hogares con educación básica o inferior. La estandarización previa asegura la

comparabilidad entre variables, y la retención de un solo componente garantiza simplicidad interpretativa. La variable resultante es `pca_educ`.

- Capacidad económica

Las variables consideradas fueron:

- `av_ipch_corr`: Promedio del ingreso per cápita del hogar a precios corrientes.
- `av_ipch_cons2023`: Promedio del ingreso per cápita del hogar a precios constantes 2023.
- `av_gpch_corr`: Promedio del gasto per cápita del hogar a precios corrientes.
- `av_gpch_cons2023`: Promedio del gasto per cápita del hogar a precios constantes 2023.
- `pbipc_corr`: PBI per cápita a precios corrientes.
- `pbipc_cons_2007`: PBI per cápita a precios constantes 2007.

La construcción del índice de capacidad económica mediante el PCA busca sintetizar de manera óptima las múltiples dimensiones del bienestar económico a nivel departamental. En indicadores estadísticos, el valor de KMO fue de 0.58, el cual supera el umbral mínimo de 0.5, lo que indica que existe suficiente varianza compartida entre las variables para proceder con el análisis factorial. Este valor es comprensible dado que el índice combina deliberadamente dos tipos de información económica: por un lado, indicadores microeconómicos de los hogares (ingreso y gasto per cápita) y por otro, indicadores macroeconómicos departamentales (PBI per cápita), considerando tanto valores nominales (corrientes) como reales (constantes).

La inclusión de varias dimensiones, valores corrientes y constantes, o datos microeconómicos y macroeconómicos, responde a la necesidad de capturar, por un lado,

tanto la capacidad económica observada en términos nominales como el poder adquisitivo real ajustado por inflación. Asimismo, el componente microeconómico y macroeconómico permite disminuir la heterogeneidad inherente a la agregación regional. Esta aproximación dual proporciona una visión más completa de la capacidad económica que el uso exclusivo de una sola de estas medidas. El scree plot, junto con la regla de Kaiser (autovalor > 1), confirma que un único componente principal explica una proporción significativa de la varianza (82.6%), indicando que el PCA logra sintetizar eficientemente estas dimensiones interrelacionadas en un índice unidimensional.

El índice resultante (*pca_capaeco*) ofrece una interpretación clara y consistente: los valores altos representan departamentos con mayor capacidad económica integral, caracterizados por altos ingresos y gastos per cápita en los hogares (tanto nominales como reales) junto con un elevado PBI per cápita. Por el contrario, los valores bajos del índice identifican departamentos con limitaciones estructurales de desarrollo, donde tanto los hogares como la economía departamental en su conjunto muestran menores niveles de ingresos y producción.

- Vulnerabilidad económica

Las variables consideradas fueron:

- *th_pov*: tasa de pobreza.
- *tp_empinf*: tasa de empleo informal.

La construcción del índice de vulnerabilidad económica mediante reflejada una proporción de varianza explicada del 81% confirma que estas métricas comparten una dimensión latente subyacente: la vulnerabilidad económica estructural. El único componente retenido (con eigenvalue de 1.62) sintetiza eficientemente esta relación. Así, estos

indicadores validan el enfoque PCA sobre alternativas como el promedio simple o índices aditivos.

La interpretación del índice se realiza en términos de lo siguiente: valores altos del índice indican mayor tasa de pobreza y empleo informal, por tanto, mayor vulnerabilidad económica sistémica. Valores bajos indican lo contrario: menor pobreza y menor informalidad laboral, asociados a economías locales más resilientes. La pobreza y el empleo informal son fenómenos interconectados en economías en desarrollo. El empleo informal (sin protección social) suele ser tanto causa como consecuencia de la pobreza, y el PCA captura esta circularidad.

Cabe señalar que, mientras *pca_capaeco* captura la capacidad productiva y de consumo a través de ingresos, gastos y PBI departamental -reflejando recursos económicos disponibles-, el *pca_vuleco* se enfoca específicamente en factores de precariedad (pobreza e informalidad laboral) que representan riesgos y limitaciones socioeconómicas. Esta diferenciación es consistente considerando que se distingue entre capacidades económicas (recursos positivos) y vulnerabilidades (factores de riesgo), donde una región podría mostrar simultáneamente alta capacidad económica, pero persistencia de bolsones de vulnerabilidad.

Transformación y Normalización de los Índices PCA a Escala 0-100

Para garantizar la comparabilidad directa con la variable dependiente (tasa de uso de internet en porcentaje) y facilitar la interpretación de los resultados, los índices generados mediante PCA fueron transformados a una escala común de 0 a 100. Esta normalización se realizó aplicando el método Min-Max, que preserva la distribución original de los datos

mientras reescala sus valores a un rango fijo. Matemáticamente, la transformación sigue la fórmula:

$$\text{Índice}_{[0,100]} = 100 \times \left(\frac{\text{Score PCA} - \text{Mínimo}}{\text{Máximo} - \text{Mínimo}} \right)$$

donde el *mínimo* y *máximo* corresponden a los valores extremos observados en cada índice antes de la transformación. Este enfoque ofrece tres ventajas clave:

- Los valores ahora representan posiciones relativas dentro de un rango porcentual, donde 0 indica el peor escenario observado en la muestra y 100 el mejor.
- Al compartir la misma escala (0-100) con la variable de uso de internet, se simplifica la comparación de coeficientes en modelos de regresión y se evitan sesgos por diferencias en las unidades de medida.
- La transformación es lineal, por lo que no altera las correlaciones entre variables ni la estructura subyacente capturada por el PCA.

La interpretación de los índices generados por PCA es clara. Sin embargo, en términos de rangos de los valores que adoptan estos abarcan desde valores negativos a positivos. Para homogenizar el dominio de análisis con la variable a modelar (tasa de uso de internet e n%) estos índices obtenidos por PCA se llevaron a una escala de 0 a 100 usando

Barreras de percepción

Para examinar las brechas en la adopción de internet entre zonas urbanas y rurales, el estudio consideró tres barreras fundamentales: (1) desconocimiento tecnológico (*nsabe*), (2) falta de interés percibido (*nniu*) y (3) limitaciones económicas (*npxn*). Inicialmente, se evaluó la posibilidad de integrar estas dimensiones mediante un Análisis de Componentes Principales (PCA). Sin embargo, el bajo valor del KMO (<0.5) indicó falta de adecuación muestral, revelando que estas variables no comparten suficiente varianza común para

justificar un índice por PCA. Esta limitación se explica porque en la ERESTEL, los encuestados seleccionaban solo una barrera principal (diseño no multirespuesta), lo que genera una estructura de datos mutuamente excluyente.

Ante esto, se implementó una estrategia dual:

- Análisis individual de cada barrera, que permite identificar patrones específicos (ej.: el desconocimiento tecnológico como obstáculo predominante en zonas rurales).
- Creación de una variable compuesta (ranouso) mediante promedio simple de las tres dimensiones, la cual - pese a sus limitaciones - ofrece una métrica sintética útil para modelos multivariados. Esta aproximación se justifica porque:
 - ✓ Captura la intensidad relativa de cada barrera a nivel departamental.
 - ✓ Es coherente con la naturaleza unidimensional de la pregunta original (donde los encuestados priorizaban un único factor).
 - ✓ Permite comparaciones territoriales integradas, complementando los análisis desagregados.

Si bien el promedio no refleja correlaciones subyacentes (como lo haría un PCA válido), su simplicidad y transparencia lo convierten en una herramienta auxiliar válida para explorar tendencias generales, reconociendo las limitaciones de los datos disponibles, pero extrayendo *insights* relevantes para políticas públicas.

Así, las variables construidas o recopiladas por cada departamento o región son presentadas en la Tabla 2 (en el Anexo 4 se presentan todas las variables trabajadas en el modelo y que finalmente se redujeron a las mostradas en la Tabla 2).

Tabla 3*Variables utilizadas en las regresiones*

Tipo	Variable	Código	Promedio	
			Urbano	Rural
endógena	Tasa de la población que usa Internet	tp_usint	70.8397	50.2603
explicativa	Infraestructura de Internet	pca_inter	75.2929	49.0983
explicativa	Servicios básicos	pca_servbas	61.6457	40.0872
explicativa	Educación	pca_educ	66.1036	26.1363
explicativa	Capacidad económica	pca_capeco	44.5123	17.3307
explicativa	Vulnerabilidad económica	pca_vuleco	34.1820	74.3352
explicativa	Calidad de la red de Internet	pca_cali	73.2480	71.2639
explicativa	Edad	pca_edad	54.8289	57.4095
explicativa	No sabe (Tasa de población que no usa Internet porque no sabe usar Internet o computadora)	nsabe	38.6884	40.1877
explicativa	No le interesa (Tasa de población que no usa Internet porque no lo necesita, no le interesa o no le ve utilidad)	nniu	42.3406	42.1952
explicativa	Muy caro (Tasa de población que no usa Internet porque es muy caro)	nvpn	7.7154	5.1174
explicativa	Razones para no uso	ranouso	29.5815	29.1668

Elaboración propia.

3.3. LIMITACIONES

Las limitaciones respecto a la metodología de descomposición empleada, ya en la sección

3.1 se ha resumido las críticas que la propia literatura ha hecho sobre la misma.

Luego, por el lado del modelo empleado, siempre está presente el posible problema de sesgo por error de especificación o de variables omitidas. La selección e inclusión de las variables económicas explicativas específicas en el modelo ha estado restringida por la disponibilidad de información para el nivel elegido de la unidad de análisis. Así por ejemplo, no ha sido posible incluir variables o indicadores que recojan directamente el estado o la calidad del servicio de Internet (como velocidad, latencia, etc.) debido a la falta de información disponible para todo el período analizado.

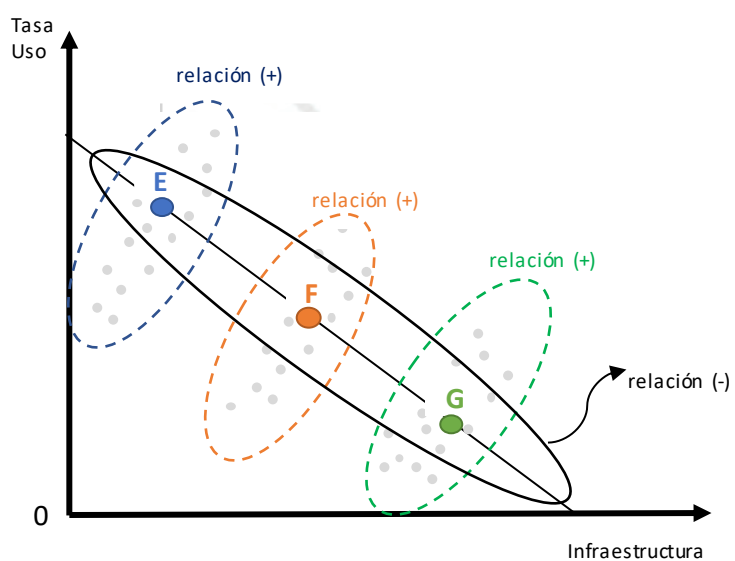
Finalmente, entre las limitantes referidas al manejo de los datos, a la parte empírica del trabajo, primero se tiene el hecho que se ha empleado un modelo *pooled* y no un modelo de panel data, aun cuando en muchos de los trabajos empíricos sobre las brechas de uso de Internet se estiman modelos de panel data. Al respecto, debe precisarse que siendo nuestra unidad de análisis las regiones del Perú, y lo que se busca es identificar variables que explican las diferencias en el uso del internet entre grupos poblacionales con menor ruralidad (población urbana) y mayor ruralidad (población rural), es posible que algunas regiones inicialmente “con mayor ruralidad” en el tiempo puedan adoptar una característica de “menor ruralidad”. Pero en el fondo este aspecto es más una sutileza que un cuestionamiento fuerte a la metodología. Además, como se verá más adelante, la agrupación de regiones como urbanas o rurales se ha mantenido estable para el período analizado.

Un segundo aspecto tiene que ver con una posible “falacia de la agregación de datos”. A nivel regional, es posible que exista una relación positiva entre infraestructura y tasa de uso de internet; pero a nivel nacional (total) al tomar los datos promedio de cada región, es posible que la relación resulte negativa. Este problema se solucionaría si se trabajara con información a nivel de centros poblados; sin embargo, mucha información necesaria

-en particular la variable dependiente- no se encuentra disponible a dicho nivel de desagregación: la tasa de uso (e incluso la tasa de acceso) de internet no se encuentra disponible a nivel de centro poblado. Precisamente, se reconoce que al trabajar con mayor capilaridad (al pasar de datos regionales a los niveles provinciales, distritales, e incluso a nivel de centro poblado), es posible que los resultados cambien.

Figura 29

Posible sesgo de agregación de variables



Luego, si bien la agregación de datos (desde el nivel de hogares) a nivel de regiones o departamentos puede esconder heterogeneidades internas dentro de cada región (dentro de la unidad de análisis), como ya fue mencionado el enfoque territorial usado compensa dicho efecto.

CAPÍTULO 4: RESULTADOS Y DISCUSIÓN

4.1. ESTIMACIONES Y RESULTADOS

Con el fin de separar la muestra en los dos grupos a ser analizados y sometidos a la metodología de descomposición Oaxaca-Blinder, se realizó un análisis de clúster para la segmentación de regiones, lo cual permitirá cuantificar cómo factores socioeconómicos, de infraestructura y preferencias explican las diferencias observadas entre regiones urbanas y rurales.

El análisis de clúster se implementó mediante el algoritmo k-means con $k=2$, utilizando como variable de segmentación la tasa de población rural (tp_rur). Esta técnica permitió dividir las regiones en dos grupos naturales: aquellas con mayor concentración de población urbana (y potencialmente con alto uso de Internet) y aquellas con mayor ruralidad (y potencialmente con bajo uso de Internet)²⁷.

La variable identificadora del grupo:

- Grupo = 0 (regiones con menor ruralidad): Urbano
- Grupo = 1 (regiones con mayor ruralidad): Rural

Así, se observa que las regiones clasificadas como urbanas (Grupo = 0) muestran tasas de acceso y uso de internet mayores a las regiones clasificadas como rurales (Grupo = 1).

²⁷ Se estableció una semilla aleatoria (set seed 12345) para garantizar la replicabilidad de los resultados.

Tabla 4*Clasificación de las regiones*

Urbano	Rural
Arequipa	Amazonas
Callao	Ancash
Ica	Apurímac
La Libertad	Ayacucho
Lambayeque	Cajamarca
Lima	Cusco
Madre de Dios	Huancavelica
Moquegua	Huánuco
Piura	Junin
Tacna	Loreto
Tumbes	Pasco
Ucayali	Puno
	San Martín

Elaboración propia.

A partir de ello, el estudio aborda el análisis de la brecha en el uso de internet entre departamentos urbanos y rurales mediante la descomposición Oaxaca-Blinder, metodología que permite desagregar las diferencias observadas en componentes explicados por características observables y factores no observados. Para la clasificación de los departamentos, se utilizó como criterio la tasa de población rural (tp_{rur}), estableciendo dos grupos consistentes a lo largo del periodo 2018-2023: aquellos con menor tasa se consideraron urbanos y aquellos con mayor tasa, rurales. Es importante destacar que, pese a trabajar con datos en panel (150 observaciones de 24 departamentos durante 6 años), el análisis se realiza mediante regresiones pooled.

Ahora bien, se verificó empíricamente que la clasificación de cada departamento se mantuvo estable durante todo el periodo de estudio. Es decir, ningún departamento cambió de categoría (de urbano a rural o viceversa) entre 2018 y 2023. Esta estabilidad

refleja que los cambios en la composición poblacional son graduales y no presentan variaciones bruscas en periodos cortos. Segundo, incluso en el caso hipotético de que algunos departamentos hubieran cambiado de categoría, esto no afectaría sustancialmente los resultados, ya que la descomposición Oaxaca-Blinder trabaja con promedios grupales y no con observaciones individuales. Es decir, lo relevante es la diferencia promedio entre los dos grupos definidos, no la trayectoria específica de cada unidad territorial.

El uso de regresiones *pooled*, en lugar de modelos de efectos fijos o aleatorios, se justifica porque el interés central del estudio no está en la variación temporal intra-departamental, sino en comparar promedios estructurales entre tipos de departamentos (urbanos vs. rurales). Además, la descomposición requiere estimar modelos separados para cada grupo (urbanos/rurales), lo que complica la implementación de efectos fijos/aleatorios. El enfoque *pooled* permite mantener la coherencia metodológica original de la técnica, que fue diseñada para datos *cross-section*. Además, Como se verificó, ningún departamento cambió de categoría grupal durante 2018-2023. Esta constancia hace que los efectos fijos departamentales se capturen implícitamente en la variable grupal (urbano/rural). Adicionalmente, se incluye variables de control que absorben parte de la heterogeneidad no observable (e.g. indicadores de desarrollo infraestructural).

Así pues, esta aproximación metodológica permite capturar de manera efectiva las diferencias persistentes entre ambos tipos de departamentos, que es precisamente el objeto de estudio de esta investigación sobre la brecha de uso de internet en Perú.

A continuación se muestran y analizan los principales resultados obtenidos con la descomposición Oaxaca-Blinder para 5 modelos analizados, los cuales cada vez se van

complejizando en cuanto a su composición, y donde el último modelo trabajado es un modelo que recoge el criterio de parsimonia. En el Anexo N° 5 se presentan los resultados en detalle, para cada uno de los modelos analizados, tanto de las estimaciones directas en cada grupo como de la descomposición Oaxaca-Blinder, correspondiente.

Resulta importante precisar que la robustez de los resultados de la metodología desarrollada se valida a nivel de cada regresión lineal para cada grupo en cada uno de los 5 modelos y no en los resultados de la descomposición Oaxaca-Blinder, pues la descomposición como tal es una construcción a partir de los resultados de las regresiones individuales. En el Anexo 5 se muestran los indicadores de robustez y grados de ajustes tales como el estadístico F (F-stat), el R cuadrado (R^2) y la Raíz del Error Cuadrático Medio (Root MSE).

Debido a que no resultan directamente interpretables los valores asociados a cada variable dentro de la descomposición Oaxaca-Blinder, resulta pertinente presentar algunas precisiones.

En primer lugar, debe notarse que las ecuaciones (4), (5) y (6) son maneras alternativas de presentar la referida descomposición. De hecho, la ecuación (6) es similar al planteamiento original de Blinder (1973) y al mismo tiempo es la representación de la descomposición que guarda relación directa con la manera como el paquete estadístico Stata muestra los resultados de la referida descomposición.

$\Delta\bar{Y}$ es la brecha total observada (diferencia total observada en la nomenclatura original) evaluada en el promedio de los grupos analizados, y las características observadas promedio (\bar{x}) son llamadas también las dotaciones (*endowments* en la nomenclatura

original). Luego el efecto dotación (que corresponde al componente D en la ecuación (4) y se asocia a la parte explicada) está representado por el componente E en la ecuación de Blinder (1973); mientras que el efecto coeficiente (componente C de la ecuación (4)) corresponde al componente C de Blinder (1973) y el efecto discriminación (puro) es recogido en el componente U de Blinder (1973) y en el componente B de la ecuación (4).

Tabla 5

Descomposición Oaxaca-Blinder: output del Stata

Blinder-Oaxaca decomposition		Number of obs	=	150		
		Model	=	linear		
Group 1: grupo = 0		N of obs 1	=	72		
Group 2: grupo = 1		N of obs 2	=	78		
explained: (X1 - X2) * b1						
unexplained: X2 * (b1 - b2)						
tp_usint	Coefficient	Robust std. err.	z	P> z	[95% conf. interval]	
overall						
group_1	70.83973	1.235212	57.35	0.000	68.41876	73.2607
group_2	50.26028	1.558004	32.26	0.000	47.20665	53.31391
difference	20.57944	1.988247	10.35	0.000	16.68255	24.47634
explained	10.94077	2.239157	4.89	0.000	6.552106	15.32944
unexplained	9.638671	1.756731	5.49	0.000	6.195541	13.0818
explained						
pca_inter_	13.74383	1.86688	7.36	0.000	10.08481	17.40284
pca_servbas_	1.798779	.743943	2.42	0.016	.3406772	3.25688
pca_educ_	-2.730516	2.194232	-1.24	0.213	-7.031133	1.5701
pca_capaeco_	5.269061	1.026325	5.13	0.000	3.257502	7.280621
pca_vuleco_	-7.292396	2.074141	-3.52	0.000	-11.35764	-3.227154
ranouso	.1520191	.1782049	0.85	0.394	-.1972561	.5012944
unexplained						
pca_inter_	-4.059015	3.372936	-1.20	0.229	-10.66985	2.551819
pca_servbas_	9.451897	2.678849	3.53	0.000	4.20145	14.70234
pca_educ_	-7.841219	2.017365	-3.89	0.000	-11.79518	-3.887257
pca_capaeco_	-5.471607	2.937581	-1.86	0.063	-11.22916	.2859452
pca_vuleco_	-14.67142	8.013756	-1.83	0.067	-30.37809	1.035257
ranouso	26.45034	11.68863	2.26	0.024	3.541048	49.35964
_cons	5.779689	14.41931	0.40	0.689	-22.48164	34.04102

$$\sum_{j=1}^K \beta_j^U (\bar{x}_j^U - \bar{x}_j^R)$$

Efecto Dotación

$$\sum_{j=1}^K (\beta_j^U - \beta_j^R) \bar{x}_j^R$$

Efecto Coeficientes

$$(\beta_0^U - \beta_0^R)$$

Efecto Discriminación

Elaboración propia.

La literatura que ha trabajado la descomposición de Oaxaca-Blinder en el mercado de las telecomunicaciones (brecha digital en sus diversas variantes, niveles de agregación conceptual y regional), al momento de analizar la parte no explicada se ha centrado básicamente en encontrar la incidencia total de dicha parte no explicada sobre la brecha digital (atribuible a características no observables de las variables y a aspectos estructurales) antes que en buscar o dar explicación o entendimiento directo de la contribución marginal (o retornos) de cada variable dentro de esta parte no explicada sobre la brecha (en el entendimiento de la brecha).

En ese sentido, en el presente trabajo se propone una explicación en torno a los retornos de cada variable sobre la brecha, haciendo la salvedad que dicha explicación tiene la característica de “posibles causas” antes que de “causas seguras”.

El estudio empleó una regresión pooled con estimación robusta de varianzas para analizar datos de los 25 departamentos peruanos observados anualmente entre 2018 y 2023, analizando las brechas en el uso de internet (tp_usint) entre zonas urbanas ($grupo1=0$) y rurales ($grupo1=1$). Esta elección metodológica corrige potenciales problemas de heterocedasticidad y correlación serial, coadyuvando en la robustez de los errores estándar y los tests de significancia.

- Primer Modelo - Modelo Base (brecha de uso explicada solo por infraestructura básica de internet).

El análisis de los resultados revela una brecha significativa en el uso de internet entre departamentos urbanos (Grupo 1) y rurales (Grupo 2) del Perú. Los resultados muestran

que la tasa promedio de uso de internet en zonas urbanas alcanza el 70.84%, mientras que en zonas rurales se reduce al 50.26%, generando una diferencia absoluta de 20.58 puntos porcentuales. Esta brecha se descompone en dos componentes principales: un 77.8% (16.00 puntos) atribuible a diferencias en características observables (parte explicada) y un 22.2% (4.57 puntos) asociado a factores no observados (parte no explicada).

La totalidad de la brecha explicada (16.00 puntos) se debe a diferencias en el índice de infraestructura de internet (*pca_inter_*), única variable utilizada en esta regresión, lo que indica que las disparidades en acceso a telefonía móvil y conexión a internet son el principal determinante de la desigualdad digital entre zonas urbanas y rurales. Este hallazgo es estadísticamente significativo y robusto, con un intervalo de confianza del 95% que va de 12.47 a 19.53 puntos porcentuales. El coeficiente negativo no significativo para *pca_inter_* en la parte no explicada (-2.38) sugiere, por ahora, que no hay evidencia suficiente para afirmar que el retorno de la infraestructura (su efecto marginal en el uso de internet) difiera significativamente entre ambos grupos.



Tabla 6

Primer Modelo. Resultados Descomposición Oaxaca-Blinder

Blinder-Oaxaca decomposition		Number of obs	=	150		
		Model	=	linear		
Group 1: grupo = 0		N of obs 1	=	72		
Group 2: grupo = 1		N of obs 2	=	78		
<p>explained: $(X1 - X2) * b1$ unexplained: $X2 * (b1 - b2)$</p>						
tp_usint	Coefficient	Robust std. err.	z	P> z	[95% conf. interval]	
overall						
group_1	70.83973	1.229946	57.60	0.000	68.42908	73.25038
group_2	50.26028	1.552107	32.38	0.000	47.21821	53.30236
difference	20.57944	1.980355	10.39	0.000	16.69802	24.46087
explained	16.00453	1.800932	8.89	0.000	12.47477	19.53429
unexplained	4.574914	1.179672	3.88	0.000	2.2628	6.887029
explained						
pca_inter_	16.00453	1.800932	8.89	0.000	12.47477	19.53429
unexplained						
pca_inter_	-2.3803	2.590307	-0.92	0.358	-7.457208	2.696608
_cons	6.955214	3.26047	2.13	0.033	.5648114	13.34562

Elaboración propia.

El componente no explicado de la brecha (4.57 puntos) alcanza significancia estadística y se asocia principalmente al intercepto (6.96 puntos), lo que refleja la existencia de factores no observados que persisten incluso al controlar por diferencias en infraestructura. Para esta primera regresión, la magnitud relativa de los componentes (77.8% explicado vs. 22.2% no explicado) sugiere que, aunque mejorar la infraestructura digital reduciría sustancialmente la brecha, sería necesario complementar estas políticas con intervenciones que aborden las barreras no observadas para lograr una equidad digital plena.

Momentáneamente, estos hallazgos sugieren importantes implicancias para políticas públicas, destacando que: 1) la expansión de infraestructura digital debe ser prioritaria, pero 2) requiere complementarse con programas de capacitación digital y asequibilidad para abordar los factores no observados que mantienen la brecha residual.

- Segundo Modelo - Modelo Intermedio (se incluyen otros factores de infraestructura, así como la edad y educación).

Esta regresión sigue revelando una brecha persistente en el uso de internet entre zonas urbanas (70.8%) y rurales (50.3%) en Perú durante el periodo 2018-2023, con una diferencia de 20.6 puntos porcentuales estadísticamente significativa ($p < 0.001$). La descomposición muestra que el 75.5% de esta brecha (15.5 puntos) se explica por diferencias en características observables, mientras que el 24.5% restante (5.0 puntos) corresponde a factores no observados. La infraestructura de internet sigue emergiendo como el principal determinante, explicando por sí sola 15.8 puntos de la brecha (76.9% del total), siendo que este efecto dominante persiste incluso al controlar por otras variables, confirmando su papel central en las desigualdades digitales.

Tabla 7

Segundo Modelo. Resultados Descomposición Oaxaca-Blinder

Blinder-Oaxaca decomposition		Number of obs	=	150		
		Model	=	linear		
Group 1: grupo = 0		N of obs 1	=	72		
Group 2: grupo = 1		N of obs 2	=	78		
explained: $(X1 - X2) * b1$ unexplained: $X2 * (b1 - b2)$						
tp_usint	Coefficient	Robust std. err.	z	P> z	[95% conf. interval]	
overall						
group_1	70.83973	1.236006	57.31	0.000	68.4172	73.26225
group_2	50.26028	1.558128	32.26	0.000	47.20641	53.31416
difference	20.57944	1.988838	10.35	0.000	16.68139	24.47749
explained	15.53048	2.216916	7.01	0.000	11.18541	19.87556
unexplained	5.048962	1.7873	2.82	0.005	1.545919	8.552005
explained						
pca_inter_	15.84589	2.090144	7.58	0.000	11.74928	19.9425
pca_cali_	-.0825045	.0924477	-0.89	0.372	-.2636987	.0986896
pca_servbas_	2.416	1.343568	1.80	0.072	-.2173439	5.049344
pca_edad_	.0720206	.1469261	0.49	0.624	-.2159492	.3599904
pca_educ_	-2.720923	1.631207	-1.67	0.095	-5.918031	.4761849
unexplained						
pca_inter_	-6.971974	3.194907	-2.18	0.029	-13.23388	-.7100718
pca_cali_	-2.432476	11.54239	-0.21	0.833	-25.05514	20.19019
pca_servbas_	16.60859	3.526099	4.71	0.000	9.69756	23.51961
pca_edad_	-5.222644	3.230208	-1.62	0.106	-11.55374	1.108448
pca_educ_	-6.169109	1.842444	-3.35	0.001	-9.780233	-2.557985
_cons	9.236578	12.42734	0.74	0.457	-15.12057	33.59372

Elaboración propia.

Un hallazgo particular es la heterogeneidad en los retornos de las inversiones entre zonas urbanas y rurales. Mientras la infraestructura muestra un efecto marginal 28% menor en áreas rurales (-6.97), los servicios básicos presentan el patrón opuesto, con un impacto significativamente mayor en estas zonas (16.61). Esto sugiere que, en contextos rurales, el acceso a servicios como agua y electricidad potencia la adopción de internet más que en áreas urbanas, posiblemente por actuar como complementos necesarios para el uso tecnológico. La educación formal revela otra asimetría preocupante: aunque contribuye a reducir la brecha en la parte explicada (-2.72), su efecto marginal es notoriamente menor en zonas rurales (-6.17), indicando que los

años de escolaridad no se traducen en habilidades digitales con la misma eficacia que en contextos urbanos.

- Tercer Modelo - Modelo Completo I (se incorporan las percepciones de los consumidores).

El modelo completo de descomposición Oaxaca-Blinder, que incorpora índices de infraestructura, capital humano, condiciones socioeconómicas y barreras percibidas, revela una brecha consolidada de 20.6 puntos porcentuales en el uso de internet entre zonas urbanas (70.8%) y rurales (50.3%). La descomposición muestra que el 61.7% de esta diferencia (12.7 puntos) se explica por características observables, mientras que el 38.3% restante (7.9 puntos) corresponde a factores no observados.

Dos nuevos hallazgos emergen con especial relevancia: la capacidad económica contribuye significativamente a reducir la brecha (5.6 puntos), mientras que la vulnerabilidad económica (*pca_vuleco_*) la amplía (-7.3 puntos). Este efecto dual de los factores económicos subraya que las desigualdades digitales reflejan profundas asimetrías en desarrollo regional. Las barreras percibidas (*nsabe*, *nui*, *npxn*) no muestran significancia estadística individual ($p > 0.68$), aunque el costo percibido (*npxn*) presenta un margen de relevancia (0.58 puntos, $p = 0.095$), sugiriendo que las políticas de asequibilidad podrían tener impacto focalizado siempre que se realicen intervenciones paralelas en otros frentes.

Tabla 8

Tercer Modelo. Resultados Descomposición Oaxaca-Blinder

Blinder-Oaxaca decomposition		Number of obs	=	150		
		Model	=	linear		
Group 1: grupo = 0		N of obs 1	=	72		
Group 2: grupo = 1		N of obs 2	=	78		
explained: $(X1 - X2) * b1$ unexplained: $X2 * (b1 - b2)$						
tp_usint	Coefficient	Robust std. err.	z	P> z	[95% conf. interval]	
overall						
group_1	70.83973	1.238801	57.18	0.000	68.41172	73.26773
group_2	50.26028	1.564059	32.13	0.000	47.19478	53.32578
difference	20.57944	1.995221	10.31	0.000	16.66888	24.49001
explained	12.70671	2.385938	5.33	0.000	8.030357	17.38306
unexplained	7.872735	1.902151	4.14	0.000	4.144587	11.60088
explained						
pca_inter_	13.15415	1.999954	6.58	0.000	9.234314	17.07399
pca_cali_	.0001819	.0423382	0.00	0.997	-.0827994	.0831631
pca_servbas_	2.534481	1.209464	2.10	0.036	.163974	4.904987
pca_edad_	.1170975	.1839993	0.64	0.525	-.2435344	.4777294
pca_educ_	-1.900278	2.282061	-0.83	0.405	-6.373035	2.572478
pca_capaeco_	5.591836	1.172202	4.77	0.000	3.294363	7.889309
pca_vuleco_	-7.31228	2.321961	-3.15	0.002	-11.86324	-2.76132
nsabe	-.0755722	.1882558	-0.40	0.688	-.4445468	.2934024
nniu	.0170011	.2308715	0.07	0.941	-.4354986	.4695009
nxpn	.5800913	.3478265	1.67	0.095	-.1016362	1.261819
unexplained						
pca_inter_	-5.322242	3.962838	-1.34	0.179	-13.08926	2.444778
pca_cali_	-6.256003	10.98047	-0.57	0.569	-27.77734	15.26533
pca_servbas_	11.73256	3.814246	3.08	0.002	4.256773	19.20834
pca_edad_	-4.243816	3.372085	-1.26	0.208	-10.85298	2.365349
pca_educ_	-7.386182	2.353354	-3.14	0.002	-11.99867	-2.773693
pca_capaeco_	-5.328414	3.257514	-1.64	0.102	-11.71302	1.056195
pca_vuleco_	-14.74601	9.413928	-1.57	0.117	-33.19697	3.704947
nsabe	9.007051	5.874797	1.53	0.125	-2.507339	20.52144
nniu	10.46066	6.237421	1.68	0.094	-1.764464	22.68578
nxpn	1.300106	1.166498	1.11	0.265	-.9861879	3.586401
_cons	18.65503	20.15544	0.93	0.355	-20.8489	58.15897

Elaboración propia.

La parte no explicada de la brecha ofrece insights valiosos sobre efectividad diferencial de políticas. Los servicios básicos tienen un retorno 34% mayor en zonas rurales (11.7 puntos), confirmando su rol como facilitadores críticos en contextos de baja conectividad. Contrariamente, la educación muestra una brecha de efectividad preocupante: su impacto en adopción tecnológica es 7.4 puntos menor en áreas rurales, evidenciando fallas en la

transferencia de habilidades digitales a través del sistema educativo formal. El intercepto (18.7 puntos, $p=0.355$), aunque ahora no significativo, sugiere la existencia de factores omitidos, que podrían requerir investigación cualitativa complementaria.

Esta tercera regresión plantearía tres implicaciones clave para política pública: primero, las intervenciones deben combinar infraestructura con programas de inclusión financiera y reducción de vulnerabilidades; segundo, el sistema educativo rural necesita reformas curriculares para cerrar la brecha en habilidades digitales; y tercero, las percepciones sobre utilidad y costo del internet requieren abordajes culturalmente adaptados. La robustez del modelo queda demostrada por la estabilidad del coeficiente de infraestructura frente a múltiples controles, y la emergencia de nuevos predictores consistentes con la literatura sobre determinantes de adopción tecnológica.

- Cuarto Modelo - Modelo Completo II (se precisa el rol clave del factor demanda por medio de una verificación a través de una medida sintetizada de barreras de percepción en el uso de internet).

El modelo completo que incorpora el índice sintético de barreras percibidas (ranouso) confirma la persistencia de una brecha significativa de 20.6 puntos porcentuales en el uso de internet entre zonas urbanas (70.8%) y rurales (50.3%). La descomposición revela que el 59.3% de esta diferencia (12.2 puntos) se explica por factores observables, mientras que el 40.7% restante (8.4 puntos) corresponde a componentes no explicados. La infraestructura digital mantiene su rol dominante, explicando 12.9 puntos (62.7% de la brecha total, aunque su contribución absoluta disminuye levemente respecto a modelos anteriores al controlar por el efecto sintetizado de las barreras percibidas.

Tabla 9

Cuarto Modelo. Resultados Descomposición Oaxaca-Blinder

Blinder-Oaxaca decomposition		Number of obs	=	150		
		Model	=	linear		
Group 1: grupo = 0		N of obs 1	=	72		
Group 2: grupo = 1		N of obs 2	=	78		
<p>explained: $(X1 - X2) * b1$ unexplained: $X2 * (b1 - b2)$</p>						
tp_usint	Coefficient	Robust std. err.	z	P> z	[95% conf. interval]	
overall						
group_1	70.83973	1.237335	57.25	0.000	68.41459	73.26486
group_2	50.26028	1.561018	32.20	0.000	47.20074	53.31982
difference	20.57944	1.991928	10.33	0.000	16.67534	24.48355
explained	12.19804	2.346227	5.20	0.000	7.599522	16.79656
unexplained	8.381402	1.883731	4.45	0.000	4.689356	12.07345
explained						
pca_inter_	12.91384	1.995023	6.47	0.000	9.003664	16.82401
pca_cali_	.0222376	.0479858	0.46	0.643	-.0718129	.1162881
pca_servbas_	3.13504	1.225421	2.56	0.011	.7332581	5.536821
pca_edad_	.1528726	.2278723	0.67	0.502	-.2937488	.5994941
pca_educ_	-1.787008	2.177823	-0.82	0.412	-6.055463	2.481447
pca_capaeco_	5.826167	1.189814	4.90	0.000	3.494175	8.158159
pca_vuleco_	-8.189913	2.082861	-3.93	0.000	-12.27225	-4.10758
ranouso	.1248099	.1601694	0.78	0.436	-.1891165	.4387362
unexplained						
pca_inter_	-6.058392	3.905897	-1.55	0.121	-13.71381	1.597025
pca_cali_	-5.100607	10.72396	-0.48	0.634	-26.11918	15.91797
pca_servbas_	13.18365	3.729129	3.54	0.000	5.874693	20.49261
pca_edad_	-4.744793	3.510627	-1.35	0.177	-11.6255	2.135909
pca_educ_	-7.298496	2.284198	-3.20	0.001	-11.77544	-2.821551
pca_capaeco_	-5.490484	3.003401	-1.83	0.068	-11.37704	.3960735
pca_vuleco_	-13.10621	8.871351	-1.48	0.140	-30.49374	4.281319
ranouso	23.11428	11.77073	1.96	0.050	.0440613	46.18449
_cons	13.88246	19.47339	0.71	0.476	-24.28468	52.0496

Elaboración propia.

Los resultados destacan dos patrones en la parte explicada: los servicios básicos emergen como segundo factor relevante (3.1 puntos), mientras la vulnerabilidad económica muestra un efecto inhibitorio sustancial (-8.2 puntos). La capacidad económica consolida su papel positivo (5.8 puntos), confirmando que el desarrollo económico regional actúa como facilitador clave para la adopción tecnológica. El índice de barreras percibidas (ranouso) no alcanza significancia estadística en la parte explicada (0.12 puntos, $p=0.436$), sugiriendo

que estas percepciones no explican diferencias intergrupales en el uso de internet cuando se controla por otros factores estructurales.

En la parte no explicada, tres hallazgos requieren atención especial: primero, el índice de barreras de percepción (ranouso) muestra un efecto marginal significativo (23.1 puntos), indicando que en zonas rurales, las percepciones negativas sobre el internet tienen un impacto desproporcionadamente mayor en la reducción del uso, efecto que no se captura plenamente en las diferencias de medias. En segundo lugar, llama más que la atención el hecho que la educación mantenga un impacto negativo importante (-7.29 puntos); y en tercer lugar, los servicios básicos se constituyen como elementos facilitadores (18.18 puntos).

Los resultados de la cuarta regresión tienen cuatro implicaciones para política pública: (1) las intervenciones en infraestructura deben priorizarse pero complementarse con programas de reducción de vulnerabilidades económicas; (2) el sistema educativo rural requiere adaptaciones curriculares para cerrar la brecha en habilidades digitales; (3) las percepciones negativas operan principalmente como amplificadores locales de la brecha, requiriendo estrategias de sensibilización comunitaria; y (4) los servicios básicos (agua, electricidad) potencian el impacto de las inversiones en conectividad. La robustez del modelo se refleja en la estabilidad de los coeficientes clave frente a especificaciones alternativas, validando la estrategia analítica de incorporar tanto factores estructurales como perceptuales en la descomposición de la brecha digital.

- Quinto Modelo - Modelo Parsimonioso (eliminación de variables estadísticamente no significativas al 10%)

En el caso del Quinto Modelo, que es a partir del cual se presentan algunas recomendaciones, la Tabla 8 muestra un resumen de los resultados de las regresiones iniciales para los grupos urbanos y rurales. La educación y las barreras de percepción resultan no significativas en su explicación sobre la tasa de uso de Internet en áreas urbanas; sin embargo, sí resultan significativas en zonas rurales. Asimismo, se aprecia que las barreras de percepción tienen una incidencia relativa importante en la explicación de la tasa de uso de Internet en áreas rurales: cuanto mayores son las percepciones negativas de la población sobre los beneficios del Internet, menor es su uso.

Tabla 10

Quinto Modelo. Resultados de las regresiones por grupos

	Promedios		Modelo 5			
	Urbano	Rural	Urbano		Rural	
			Coef.	Signific. 5%	Coef.	Signific. 5%
Estadísticos de ajuste						
Prueba F-Stat			74.720	0.000	128.040	0.000
R-squared			0.880		0.872	
Root MSE			3.782		5.101	
Infraestructura Internet	75.293	49.098	0.525	0.000	0.607	0.000
Servicios básicos	61.646	40.087	0.083	0.014	-0.152	0.010
Educación	66.104	26.136	-0.068	0.216	0.232	0.000
Capacidad económica	44.512	17.331	0.194	0.000	0.510	0.003
Vulnerabilidad económica	34.182	74.335	0.182	0.001	0.379	0.000
Razones para no uso	29.581	29.167	0.367	0.182	-0.540	0.071
Intercepto			5.026	0.501	-0.753	0.952

Nota: Las columnas "Coef." hacen referencia a los valores estimados de los parámetros para cada grupo; y las columnas "Signific. 5%" muestran el valor de la significancia estadística individual de cada parámetro al 5%.

Elaboración propia.

A modo de parsimonia, se ha trabajado un quinto modelo que conserva solo las variables estadísticamente significativas al 10%. Los resultados, confirma la existencia de una brecha sostenida de 20.6 puntos porcentuales en el uso de internet entre zonas urbanas (70.8%) y rurales (50.3%). La descomposición muestra que el 53.2% de esta diferencia (10.9 puntos) se atribuye a factores observables, mientras que el 46.8% restante (9.6 puntos) corresponde a componentes no explicados. La infraestructura digital se mantiene como el principal determinante, explicando 13.7 puntos (66.8% de la brecha total), con una contribución incluso mayor que en el modelo completo, lo que refuerza su carácter estructural en las desigualdades tecnológicas.

En la parte explicada, los servicios básicos consolidan su rol como segundo factor relevante (1.8 puntos), destacando que el acceso a agua y electricidad potencia la adopción de internet. La capacidad económica mantiene su efecto positivo (5.3 puntos), mientras que la vulnerabilidad económica continúa ampliando la brecha (-7.3 puntos). Estos resultados subrayan que las políticas de conectividad deben integrarse con estrategias de desarrollo económico territorial. El índice de barreras percibidas (ranouso) no alcanza significancia estadística en la parte explicada (0.15 puntos, $p=0.394$), corroborando que estas percepciones no explican diferencias intergrupales cuando se controla por otros factores estructurales.

Tabla 11

Quinto Modelo. Resultados Descomposición Oaxaca-Blinder

Blinder-Oaxaca decomposition		Number of obs	=	150		
Group 1: grupo = 0		Model	=	linear		
Group 2: grupo = 1		N of obs 1	=	72		
		N of obs 2	=	78		
explained: (X1 - X2) * b1						
unexplained: X2 * (b1 - b2)						
tp_usint	Coefficient	Robust std. err.	z	P> z	[95% conf. interval]	
overall						
group_1	70.83973	1.235212	57.35	0.000	68.41876	73.2607
group_2	50.26028	1.558004	32.26	0.000	47.20665	53.31391
difference	20.57944	1.988247	10.35	0.000	16.68255	24.47634
explained	10.94077	2.239157	4.89	0.000	6.552106	15.32944
unexplained	9.638671	1.756731	5.49	0.000	6.195541	13.0818
explained						
pca_inter_	13.74383	1.86688	7.36	0.000	10.08481	17.40284
pca_servbas_	1.798779	.743943	2.42	0.016	.3406772	3.25688
pca_educ_	-2.730516	2.194232	-1.24	0.213	-7.031133	1.5701
pca_capaeco_	5.269061	1.026325	5.13	0.000	3.257502	7.280621
pca_vuleco_	-7.292396	2.074141	-3.52	0.000	-11.35764	-3.227154
ranouso	.1520191	.1782049	0.85	0.394	-.1972561	.5012944
unexplained						
pca_inter_	-4.059015	3.372936	-1.20	0.229	-10.66985	2.551819
pca_servbas_	9.451897	2.678849	3.53	0.000	4.20145	14.70234
pca_educ_	-7.841219	2.017365	-3.89	0.000	-11.79518	-3.887257
pca_capaeco_	-5.471607	2.937581	-1.86	0.063	-11.22916	.2859452
pca_vuleco_	-14.67142	8.013756	-1.83	0.067	-30.37809	1.035257
ranouso	26.45034	11.68863	2.26	0.024	3.541048	49.35964
_cons	5.779689	14.41931	0.40	0.689	-22.48164	34.04102

Elaboración propia.

Para la parte no explicada, el modelo revela tres patrones clave: (1) los servicios básicos muestran un retorno 22% mayor en zonas rurales (9.5 puntos), confirmando su papel como facilitadores críticos en contextos de baja conectividad; (2) la educación presenta una brecha de efectividad preocupante (-7.8 puntos), evidenciando que los años de escolaridad no se traducen en adopción tecnológica con la misma eficacia que en áreas urbanas; y (3) las barreras percibidas tienen un impacto desproporcionado en zonas rurales (26.5 puntos), actuando como amplificadores locales de la brecha. Este último hallazgo sugiere que las

percepciones negativas si tienen un papel activo, afectando más intensamente el uso efectivo de internet en contextos rurales incluso cuando la infraestructura está disponible.

Los resultados del quinto modelo, plantean cuatro directrices: primero, las inversiones en infraestructura digital deben complementarse con programas de electrificación y saneamiento básico; segundo, el sistema educativo rural requiere adaptaciones curriculares específicas para desarrollar habilidades digitales prácticas; tercero, las intervenciones deben abordar las percepciones negativas mediante campañas de sensibilización comunitaria; y cuarto, las estrategias de inclusión digital deben integrarse con políticas de reducción de vulnerabilidades económicas. La consistencia de estos hallazgos frente a múltiples especificaciones modelares valida su robustez y relevancia para el diseño de políticas públicas orientadas a cerrar la brecha digital en Perú.

4.2. DISCUSIÓN DE RESULTADOS Y POSIBLES EXTENSIONES

Por lo visto en la sección previa, efectivamente, el componente no explicado puede reflejar varias dinámicas simultáneas:

- Diferencias en los “retornos” o efectos marginales que las mismas características tienen en contextos urbanos vs. rurales.
- Variables relevantes omitidas del modelo, por falta de datos o dificultad de cuantificación, fenómenos más subjetivos o difíciles de operacionalizar. Por ejemplo: barreras cognitivas, confianza en la tecnología, normas sociales, culturales, características del ecosistema digital local (disponibilidad o calidad específica del servicio de acceso a internet en las zonas rurales) o dinámicas institucionales y económicas (fallas o corrupción en la adjudicación de proyectos, relevancia percibida del internet para la vida personal, familiar y económica, incentivos y fallas de mercado).

Por eso, se sugiere, que la tesis presente los resultados como una primera aproximación cuantitativa, que ayuda a acotar el problema y orientar políticas, pero que aún no agota el análisis causal (ya que, hasta donde se ha encontrado, 'nadie' o muy poca gente ha analizado empíricamente el tema en el Perú, a ese nivel).

Las Tablas 12 y 13 muestra un resumen de la descomposición Oaxaca-Blinder para los cinco modelos. El análisis de los cinco modelos de descomposición Oaxaca-Blinder revela patrones consistentes en la brecha de uso de internet entre zonas urbanas y rurales del Perú. La diferencia promedio de 20.6 puntos porcentuales se descompone sistemáticamente en dos componentes principales: entre 53% y 78% de esta brecha se explica por diferencias en características observables, mientras que entre 22% y 47% corresponde a factores no observados. La infraestructura digital emerge como el determinante más relevante, explicando entre 62% y 77% de la brecha total en todos los modelos especificados. Sin embargo, los resultados demuestran claramente que las políticas de oferta, aunque fundamentales, no son suficientes por sí solas para cerrar la brecha digital.

Tabla 12

Resumen de la Descomposición Oaxaca-Blinder para los Modelos 1, 2 y 3

Variables	Modelo 1		Modelo 2		Modelo 3	
	Explicada	No explicada	Explicada	No explicada	Explicada	No explicada
Contribución absoluta	16.005	4.579	15.530	5.049	12.707	7.873
Contribución relativa	78%	22%	75%	25%	62%	38%
Infraestructura internet	16.005	-2.380	15.846	-6.972	13.154	-5.322
Calidad de la red			-0.083	-2.432	0.000	-6.256
Servicios básicos			2.416	16.609	2.534	11.733
Educación			-2.721	-6.169	-1.900	-7.386
Capacidad económica					5.592	-5.328
Vulnerabilidad económica					-7.312	-14.746
Edad			0.072	-5.223	0.117	-4.244
Barreras percibidas (no uso)						
No sabe					-0.076	9.007
No lo necesita, no le interesa					0.017	10.461
Muy caro					0.580	1.300
Constante		6.955		9.237		18.655

Nota:

Los valores en las columnas "Explicada" coinciden con los valores de las secciones "Explained" en los *output* del Stata cuya notación es $(X^U - X^R) * b_U$. Estos valores corresponden al componente Dotación (componente "Explicado" en la ecuación 6, componente D en la ecuación 4, y componente E en la notación de Blinder).

Los valores en las columnas "No explicada" coinciden con los valores de las secciones "Unexplained" en los *output* del Stata cuya notación es $X^R * (b_U - b_R)$. Estos valores corresponden al componente "No explicado" en la ecuación 6, que contiene 2 partes: efecto Coeficientes (componente C en la ecuación 4 y componente C en la notación Blinder) y efecto Discriminación (componente B en la ecuación 6, componente U en la notación Blinder) que se recoge en el valor de la constante estimada de la parte "Unexplained".

Elaboración propia.

Los principales determinantes observables de la brecha incluyen tres factores clave que interactúan entre sí. En primer lugar, la infraestructura insuficiente en zonas rurales, particularmente en acceso a telefonía móvil e internet fijo. En segundo lugar, las limitaciones económicas estructurales, representadas por bajo ingreso per cápita y alta informalidad laboral. Y en tercer lugar, las brechas en capital humano, donde los años de escolaridad muestran un retorno significativamente menor en el desarrollo de habilidades digitales en áreas rurales comparado con zonas urbanas.

Esta interacción se manifiesta claramente en el efecto limitado de la infraestructura en zonas rurales, donde los coeficientes entre -4.1 y -6.9 en la parte no explicada indican que su impacto depende críticamente de otros factores complementarios como habilidades digitales y reducción de costos percibidos.

Tabla 13

Resumen de la Descomposición Oaxaca-Blinder para los Modelos 4 y 5

Variables	Modelo 4		Modelo 5	
	Explicada	No explicada	Explicada	No explicada
Contribución absoluta	12.198	8.381	10.941	9.639
Contribución relativa	59%	41%	53%	47%
Infraestructura internet	12.914	-6.058	13.744	-4.059
Calidad de la red	0.022	-5.101		
Servicios básicos	3.135	13.184	1.799	9.452
Educación	-1.787	-7.298	-2.731	-7.841
Capacidad económica	5.826	-5.490	5.269	-5.472
Vulnerabilidad económica	-8.190	-13.106	-7.293	-14.671
Edad	0.153	-4.745		
Barreras percibidas (no uso)	0.125	23.114	0.152	26.450
No sabe				
No lo necesita, no le interesa				
Muy caro				
Constante		13.882		5.780

Nota:

Los valores en las columnas "Explicada" coinciden con los valores de las secciones "Explained" en los *output* del Stata cuya notación es $(X^U - X^R) * b_U$. Estos valores corresponden al componente Dotación (componente "Explicado" en la ecuación 6, componente D en la ecuación 4, y componente E en la notación de Blinder).

Los valores en las columnas "No explicada" coinciden con los valores de las secciones "Unexplained" en los *output* del Stata cuya notación es $X^R * (b_U - b_R)$. Estos valores corresponden al componente "No explicado" en la ecuación 6, que contiene 2 partes: efecto Coeficientes (componente C en la ecuación 4 y componente C en la notación Blinder) y efecto Discriminación (componente B en la ecuación 6, componente U en la notación Blinder) que se recoge en el valor de la constante estimada de la parte "Unexplained".
Elaboración propia.

La porción no explicada de la brecha (22% - 47%) sugiere la presencia de barreras no observadas que requieren atención específica. Entre estas podrían destacar fallas de

mercado como calidad desigual del servicio en zonas rurales, limitaciones institucionales en la implementación de políticas digitales a nivel subnacional, barreras culturales de desconfianza hacia entornos digitales, y externalidades negativas por efectos de red limitados.

Por su parte, la educación tiene un impacto significativamente menor en zonas rurales (-7.3 a -7.8 puntos), señalando deficiencias en la formación de habilidades digitales efectivas. Por tanto, se ha verificado que existe una importante brecha de retornos educativos en zonas rurales, dado que en el modelo 5, el índice educativo muestra un retorno 7.8 puntos menor en zonas rurales que en urbanas, evidenciando que mayores años de escolaridad no se traducen en un uso efectivo de Internet en entornos rurales. Este hallazgo valida la hipótesis 1.

Particularmente revelador es el rol de las barreras de demanda, donde el índice de percepción muestra un efecto marginal significativo en áreas rurales (23.1 - 26.5 puntos), demostrando que las actitudes negativas amplifican la brecha incluso cuando existe acceso físico. Así, las barreras perceptivas por el lado de la demanda, efectivamente se constituyen en amplificadores de la brecha, verificándose la hipótesis 2. En las regresiones 4 y 5, el índice de barreras percibidas no es significativo en la parte explicada, pero muestra un efecto no explicado elevado y estadísticamente significativo en zonas rurales (hasta 26.5 puntos). En términos prácticos, este resultado estaría evidenciando que aun cuando se cuente con acceso físico a Internet, las percepciones negativas limitan su adopción efectiva, especialmente en entornos donde el conocimiento digital es bajo.

Los resultados de las regresiones respaldan que, si bien las políticas de oferta (infraestructura, acceso tecnológico) son fundamentales para reducir la brecha, no son suficientes por sí solas. La evidencia clave que sustenta esta conclusión incluye:

- La persistencia de una brecha no explicada (22% - 47%) incluso después de controlar por infraestructura y factores económicos, lo que sugiere que existen barreras no resueltas por el mero acceso físico.
- El efecto limitado de la infraestructura en zonas rurales:
 - En la parte no explicada, el retorno de la infraestructura es significativamente menor en áreas rurales (coeficientes entre -4.1 y -6.9), indicando que su impacto depende de otros factores complementarios.
 - Esto revela que, sin intervenciones paralelas en habilidades digitales o reducción de costos percibidos, ampliar la cobertura tendrá rendimientos decrecientes.
- El rol creciente de las barreras de demanda en modelos completos:
 - El Índice de percepción muestra un efecto marginal significativo en áreas rurales (23.1 - 26.5 puntos), demostrando que las actitudes negativas amplifican la brecha cuando ya existe acceso.
 - Variables como educación tienen menor impacto en zonas rurales (-7.3 a -7.8 puntos), señalando que el sistema educativo actual no está formando habilidades digitales efectivas.

La evidencia empírica respalda que la sinergia oferta-demanda maximiza el impacto de las intervenciones. En modelos que combinan la infraestructura de Internet y las barreras de percepción, la brecha no explicada disminuye de 47% a 38%, sugiriendo que, cómo abordando ambos frentes se captura factores inicialmente "invisibles". Por tanto, se confirma la hipótesis 3 (la brecha digital como un fenómeno multidimensional) pues las

distintas regresiones muestran la importancia conjunta de la infraestructura, la capacidad económica, y las barreras perceptivas. Estos factores se activan en distintos niveles de la cadena de adopción tecnológica, evidenciando que cerrar la brecha digital requiere intervenciones simultáneas en oferta, acceso económico y uso efectivo.

Para una implementación efectiva, se recomienda: 1) pilotos territoriales integrados en departamentos con alta vulnerabilidad, pero potencial productivo (e.g. Cajamarca, Huánuco), 2) reformas curriculares para desarrollar habilidades digitales prácticas, y 3) sistemas de monitoreo que midan calidad de servicio y uso efectivo, no solo cobertura. Este enfoque multidimensional, que articula infraestructura con capacidades humanas y económicas, ofrece el camino más efectivo para cerrar sosteniblemente la brecha digital en el Perú.

Los resultados encontrados son consistentes con lo obtenido por Barrantes, Agüero y Matos (2019) en el sentido que en Perú (y en Guatemala,) la mayor parte de la brecha digital (de género) no se puede explicar por las variables observadas y, en ese sentido, la brecha no explicada se asocia con factores estructurales como normas sociales o estereotipos de género. Incluso, para algunos segmentos de la distribución del índice TIC utilizado (ICT Gender Gap – IGG) la brecha no explicada alcanza más del 70%, “lo cual se considera crítico para las políticas públicas”²⁸.

Sujeto al enfoque metodológico y los datos disponibles usados, los hallazgos de esta investigación constituyen el primer ejercicio sistemático de aplicación de la descomposición Oaxaca-Blinder al análisis de las brechas digitales en el Perú. Por lo que

²⁸ Los autores también encuentran que en países con mayor diversidad étnica y lingüística (como Perú y Guatemala), los efectos de la zona rural y del uso de lengua nativa tienden a revertirse en comparación con el resto.

los resultados deben interpretarse como: (1) una línea base para futuras investigaciones que puedan incorporar variables adicionales (e.g. competencia en mercados locales, etc.) en función al desarrollo de la información disponible, (2) un punto de partida para análisis cualitativos que exploren en profundidad los factores no observados identificados (como las barreras culturales o institucionales), y (3) una referencia para estudios posteriores que busquen medir el impacto de políticas específicas de conectividad rural.

Posibles extensiones

Este trabajo abre nuevas líneas de indagación que podrían enriquecerse con: datos a nivel distrital o centros poblados, información sobre encuestas especializadas sobre habilidades digitales. La robustez de los hallazgos principales sugiere que, pese a estas limitaciones, el estudio provee evidencia valiosa para el diseño de políticas públicas en un contexto donde escaseaban análisis cuantitativos sobre la problemática estudiada.

Por ello, siempre dependiendo de la información disponible, se puede ampliar la metodología con estimaciones del tipo panel data, para incluir análisis dinámico, es decir para analizar cómo se da la curva de "uso" una vez que el servicio de acceso a internet llega a un lugar.

En esa línea, también podría ser interesante incluir el año 2020 como punto de comparación para ver cómo evolucionó la brecha de uso del servicio de acceso a internet con la pandemia (a priori se puede decir que el uso o la demanda por el servicio de acceso a internet se incrementó durante la pandemia).

También se podrían complementar y/o contrastar los resultados obtenidos con la aplicación de la metodología de emparejamiento (*matching*), siguiendo la propuesta

desarrollada por Ñopo (2008-ii) y aplicada por Barrantes et al. (2019), como una estimación no paramétrica alternativa a la descomposición Oaxaca-Blinder.



CONCLUSIONES

El presente trabajo busca cubrir un vacío en la literatura respecto del análisis de descomposición paramétrica de la brecha digital urbano-rural, a fin de encontrar y entender, en particular, los factores no observables directamente y que sin embargo, ejercen influencia tanto en la manera como en la frecuencia en que la población rural se interrelaciona y utiliza el servicio de Internet.

Sujeto al enfoque metodológico y los datos disponibles usados, los hallazgos de esta investigación constituyen el primer ejercicio sistemático de aplicación de la descomposición Oaxaca-Blinder al análisis de las brechas digitales en el Perú. Por ese motivo, los resultados obtenidos deben interpretarse como: (1) una línea base para futuras investigaciones que puedan incorporar variables adicionales (e.g. competencia en mercados locales, etc.) en función al desarrollo de la información disponible, (2) un punto de partida para análisis cualitativos que exploren en profundidad los factores no observados identificados (como las barreras culturales o institucionales), y (3) una referencia para estudios posteriores que busquen medir el impacto de políticas específicas de conectividad rural.

En ese sentido, siendo que el componente no explicado puede reflejar varias dinámicas simultáneas:

- Diferencias en los retornos o efectos marginales que las mismas características tienen en contextos urbanos vs. rurales.
- Variables relevantes omitidas del modelo, por falta de datos o dificultad de cuantificación, fenómenos más subjetivos o difíciles de operacionalizar. Por ejemplo: barreras cognitivas, confianza en la tecnología, normas sociales, culturales,

características del ecosistema digital local (disponibilidad o calidad específica del servicio de acceso a internet en las zonas rurales) o dinámicas institucionales y económicas (fallas o corrupción en la adjudicación de proyectos, relevancia percibida del internet para la vida personal, familiar y económica, incentivos y fallas de mercado).

Los resultados obtenidos deben ser tomados como una primera aproximación cuantitativa, que ayuda a acotar el problema y orientar políticas, pero que aún no agota el análisis causal (ya que, hasta donde se ha encontrado, 'nadie' o muy poca gente ha analizado empíricamente el tema en el Perú, a ese nivel).

Una parte significativa de la brecha de uso urbano-rural no se puede explicar por las variables observadas, por lo que esta brecha no explicada se asocia con factores estructurales (no observables) persistentes más allá de la infraestructura, como lo serían las normas sociales, las percepciones de la población rural e incluso limitaciones institucionales.

Se ha evidenciado el rol estructural de los Factores No Observables, pues la cuarta y quinta regresión muestran que entre el 40% y 47% de la brecha total en el uso de Internet se explica por componentes no observados, siendo este efecto altamente significativo. Estos resultados sugieren que, incluso cuando se controla por infraestructura y variables socioeconómicas, subsisten diferencias estructurales que perpetúan la desigualdad digital.

A lo largo de las regresiones, el porcentaje de la brecha explicada por la infraestructura disminuye sistemáticamente al incluir variables de capital humano (educación), percepción y vulnerabilidad económica. Si bien la infraestructura conserva su efecto

significativo, su peso relativo cae de más del 75% a poco más del 60% en los modelos más completos, reflejando la necesidad de políticas integrales. Por tanto, se ha verificado la necesidad de preponderar el rol de la demanda en el diseño de políticas.

Esto indicaría que las políticas centradas únicamente en el despliegue físico de conectividad tienen retornos marginales decrecientes si no se acompañan de intervenciones orientadas a fortalecer capacidades de uso²⁹. Las políticas centradas exclusivamente en la oferta de infraestructura (redes, cobertura) son una condición necesaria, pero no suficiente para cerrar la brecha digital. La efectividad de estas inversiones se multiplica cuando se implementan simultáneamente programas de alfabetización digital, reducción de pobreza y sensibilización comunitaria. Withacre et al (2015) llega a una conclusión similar cuando dice que aunque la provisión de infraestructura es crucial, se debe poner más énfasis en fomentar la demanda de banda ancha en áreas rurales a través de programas de alfabetización digital, subsidios para la propiedad de computadoras y servicios de banda ancha.

El trabajo desarrollado ha permitido verificar que existe una importante brecha de retornos educativos en zonas rurales, pues en el modelo 5, el índice educativo muestra un retorno 7.8 puntos menor en zonas rurales que en urbanas, evidenciando que mayores años de escolaridad no se traducen en un uso efectivo de Internet en entornos rurales. Este hallazgo apunta a una deficiencia cualitativa en la educación rural respecto a habilidades tecnológicas. Esto permite validar la hipótesis 1.

²⁹ En modelos más completos, se observa que el impacto marginal de la infraestructura aumenta cuando se combina con servicios básicos, capital humano y medidas para mitigar la percepción negativa sobre el uso del Internet. La interacción entre factores estructurales y perceptivos valida la necesidad de un enfoque de política multinivel e intersectorial.

Las Barreras Perceptivas por el lado de la demanda, efectivamente se constituyen en amplificadores de la brecha, verificándose la hipótesis 2. En las regresiones 4 y 5, el índice de barreras percibidas no es significativo en la parte explicada, pero muestra un efecto no explicado elevado y estadísticamente significativo en zonas rurales (hasta 26.5 puntos). Esto indica que, aun con acceso físico a Internet, las percepciones negativas limitan su adopción efectiva, especialmente en entornos donde el conocimiento digital es bajo.

Se confirma la hipótesis 3 (la brecha digital como un fenómeno multidimensional) pues las distintas regresiones muestran la importancia conjunta de la infraestructura, la capacidad económica, y las barreras perceptivas. Estos factores se activan en distintos niveles de la cadena de adopción tecnológica, evidenciando que cerrar la brecha digital requiere intervenciones simultáneas en oferta, acceso económico y uso efectivo.

Este trabajo ha evidenciado la necesidad de generar información de calidad a nivel distrital o centros poblados, información sobre encuestas especializadas sobre habilidades digitales.

RECOMENDACIONES

Siendo que la presente tesis busca cubrir un vacío en la literatura respecto a la cuantificación de los aspectos no observables que inciden en la brecha de uso de Internet entre áreas urbanas y rurales del Perú, las primeras recomendaciones se orientan a la práctica profesional.

Futuras investigaciones sobre el tema tratado, podrían extender el análisis de la descomposición hacia niveles más desagregados de la unidad de análisis (provincias, distritos, centros poblados, hogares e incluso individuos) siempre que exista información disponible necesaria (algo que actualmente es restrictivo). Si bien se deberían mantener las categorías de las variables indicadas en la ecuación (1), las variables explicativas propiamente seleccionadas, así como la propia variable dependiente, podrían cambiar e incluso ser medidas de manera distinta.

Evidentemente, la disponibilidad de información condiciona también la selección de variables a ser consideradas en los modelos, y ello puede tener una incidencia en posibles sesgos por errores de especificación o variables omitidas. Por ello, urge la necesidad de generar y compilar información a diferentes niveles de capilaridad geográfica y social, y para ventanas temporales más amplias.

Lo anterior puede contribuir con el uso de otras técnicas de estimación como serían los modelos de panel data para, sobre la base de los resultados de dichas regresiones, luego aplicar las técnicas de descomposición ya sea paramétricas (a la Oaxaca-Blinder) como no paramétricas (método de emparejamientos o *matching*).

Un segundo grupo de recomendaciones tiene que ver con la formulación de políticas a nivel de gobierno que permitan abordar sobre todos aspectos estructurales por el lado de la demanda. Así, sobre la base de los resultados obtenidos (con la verificación también de las hipótesis de trabajo) se plantean algunas líneas de acción a nivel del gobierno para impulsar medidas concretas y sustentadas que contribuyan con el cierre de la brecha de uso.

Las recomendaciones de política pública deben adoptar un enfoque integral que combine estratégicamente intervenciones de oferta y demanda. En el ámbito de la oferta, de manera prioritaria pero no exclusiva, se requiere continuar y mejorar la aplicación de esquemas referidos a:

- 1) Expansión acelerada de infraestructura con estándares de calidad garantizados.
- 2) Promoción efectiva a la inversión mediante modelos público-privados,
- 3) Integración de las iniciativas de expansión de la infraestructura de Internet con servicios básicos como electrificación rural.

Estos esfuerzos deben complementarse necesariamente con políticas de demanda como la implementación o mejoras de medidas referidas a:

- 1) Programas de alfabetización digital contextualizados para la formación docente. El beneficio que obtengan los individuos por usar el Internet depende que sean conscientes del ¿“para qué” usarlo?. En este punto las habilidades digitales son importantes, se debe atender los menores niveles de educación y de habilidades digitales para así expandir las redes sociales o para buscar efectivamente información en el Internet (Tsetsi et al., 2017).

- 2) Estrategias (subsidios focalizados) orientadas a facilitar la adquisición de dispositivos que permitan el uso de Internet y el consumo del mismo (ya sea vía políticas tarifarias orientadas o esquemas de subsidio a la demanda).
- 3) Campañas de sensibilización práctica, con demostraciones comunitarias de aplicaciones agrícolas y sanitarias, por ejemplo, de manera que la población experimente y conozca de primera fuente, directamente, la utilidad o beneficio que obtiene por usar el Internet.

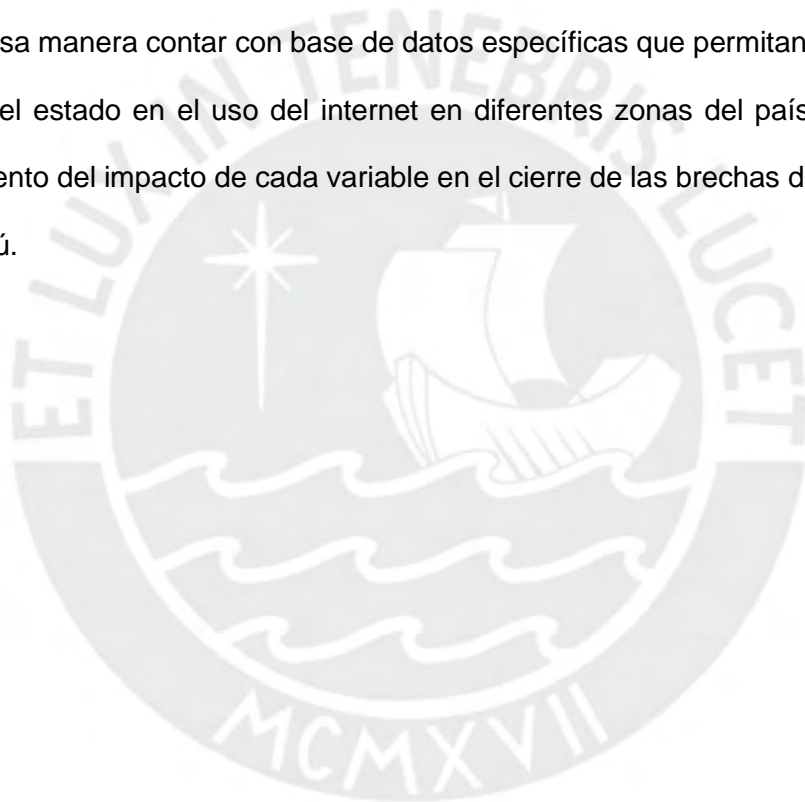
Si bien existen esfuerzos del Estado en materia de alfabetización digital (MINEDU, MTC/PRONATEL, Gobiernos regionales y locales), los resultados de la tesis sugieren que estos aún no tienen el alcance ni la profundidad necesaria para cerrar las brechas identificadas; y por tanto, estos esfuerzos deben fortalecerse mediante:

- Programas más territorializados y culturalmente sensibles.
- Intervenciones sostenidas en el tiempo, no solo talleres puntuales.
- Articulación interinstitucional efectiva entre educación, telecomunicaciones, desarrollo social y gobiernos subnacionales.
- Mecanismos de subsidio a la demanda que aborden tanto las barreras económicas como las cognitivas y culturales.

En esa línea, y siguiendo a Galperin y Arcidiacono (2021), las políticas de inclusión digital deben integrarse con estrategias para reducir las desigualdades laborales, promoviendo la capacitación digital para mujeres de bajos ingresos. Recomendaciones similares plantean Gamboa y García (2024) al puntualizar la necesidad de promover el desarrollo de habilidades digitales, crear contenidos educativos accesibles para todos los públicos, garantizar la apropiación digital, lograr la transformación digital en la educación, concretar

el fortalecimiento de las plataformas educativas y hacer efectiva la inclusión de competencias digitales en el currículo escolar.

Por ello, considerando que las recomendaciones de política apuntan a intervenciones con un nivel de capilaridad que puede llegar a nivel de centros poblados, es decir las intervenciones podrían ser diferenciadas a nivel de (tipos de) centros poblados, se considera necesario que el Estado despliegue esfuerzos para coleccionar información detallada sobre el uso del internet (ya sea a través de encuestas u otras herramientas), para de esa manera contar con base de datos específicas que permitan un conocimiento preciso del estado en el uso del internet en diferentes zonas del país, y por tanto, un conocimiento del impacto de cada variable en el cierre de las brechas de uso del internet en el Perú.



REFERENCIAS BIBLIOGRÁFICAS

- Alliance for Affordable Internet (2021). The Costs of Exclusion: Economic Consequences of the Digital Gender Gap. Web Foundation. <https://a4ai.org/wp-content/uploads/2022/03/CoE-Report-English.pdf>
- Analysys Mason. (2017). Expansión de la banda ancha móvil. Caracas: CAF. Retrieved from <https://scioteca.caf.com/handle/123456789/1084>
- Barrantes, R. (2003). Análisis de la demanda por TICs: ¿Qué es y cómo medir la pobreza digital?. Instituto de Estudios Peruanos. DIRSI. IDRC-CRDI. <https://repositorio.iep.org.pe/server/api/core/bitstreams/0b2398ba-4f8b-468e-adea-bb4981721d07/content>
- Barrantes, R., Agüero A., & Matos P. (2019). Decomposing the ICT use gender Gap for five Latin American countries. In Presented at the 13th CPR LATAM conference, Cordoba, Argentina. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3142380
- Blinder, A. (1973). Wage Discrimination: Reduced Form and Structural Estimates. *The Journal of Human Resources*, 8 (4), 436-455. <https://www.jstor.org/stable/144855>
- Bonifaz, J.L. (2024). Evaluación de impacto de los proyectos de telecomunicaciones, relacionados con el acceso a internet de banda ancha en Perú. PROINVERSIÓN. Documento de Trabajo N° 3. <https://www.investinperu.pe/es/pi/publicaciones-digitales/documentos-de-trabajo/impacto-de-los-proyectos-de-telecomunicaciones>

Bruno, G., Diglio, A., Piccolo, C., & Pipicelli, E. (2023). A reduced Composite Indicator for Digital Divide measurement at the regional level: An application to the Digital Economy and Society Index (DESI). *Technological Forecasting & Social Change*, 190, 122461. <https://www.sciencedirect.com/science/article/pii/S0040162523001464>

Castillo, C., & Chahuara, P. (2020). Análisis del estado de la demanda de los servicios públicos de telecomunicaciones 2015-2019. OSIPTEL, Dirección de Políticas Regulatorias y Competencia - Sub Dirección de Análisis Regulatorio. Documento de trabajo N° 47. <https://repositorio.osiptel.gob.pe/bitstream/handle/20.500.12630/742/DT%2047%20Informe%20Caracterizaci%3b3n%20Demanda%20SSPPTT%202015-19.pdf?sequence=1&isAllowed=y>

Chinn, M., & Fairlie, R. (2006). ICT Use in the Developing World: An Analysis of Differences in Computer and Internet Penetration. *NBER Working Paper* No. 12382. JEL No. O30, L96. <http://www.nber.org/papers/w12382>

Corrocher, N., & Ordanini, A. (2002). Measuring the digital divide: A framework for the Analysis of Cross- Country Differences. IN *Journal of Information Technology*, 17, 9-19. https://www.researchgate.net/publication/233587212_Measuring_the_Digital_Divide_A_Framework_for_the_Analysis_of_Cross-Country_Differences

De Clercq, M., D'Haese, M., & Buysse, J. (2023). Economic growth and broadband access: The European urban-rural digital divide. *Telecommunications Policy*, 47(6). Elsevier. <https://www.sciencedirect.com/science/article/pii/S0308596123000903>

DiMaggio, P., & Garip, F. (2012). Network effects and social inequality. *Annual Review of Sociology*, 38, 93–118.

https://scholar.harvard.edu/files/garip/files/dimaggio_garip_2012.pdf

Doyar, B., Rzali, S., & Dikkaya, M. (2023). Development of the ICT sector and the determinants of Internet use in the Southern Caucasus. *Telecommunications Policy*, 47, 102455.

Fairlie, R. (1999). The absence of African-American owned business: an analysis of the dynamics of selfemployment. *Journal of Labor Economics*, 17, 80–108.

Fairlie, R. (2004). Race and the digital divide. Santa Cruz Department of Economics, Working Paper Series.

<https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=5a98a819c9b0283746064bfa0aa308e02a4fa98c>

Fairlie, R. (2006). An Extension of the Blinder-Oaxaca Decomposition Technique to Logit and Probit Models. *IZA Discussion Paper Series No. 1917*. Institute for the Study of Labor.

<https://docs.iza.org/dp1917.pdf>

Federal Communication Commission (1996). In the matter of Federal-State Joint board of Universal Service, CC. Docket. No.96-45.

Galperin, H., & Arcidiacono, M. (2021). Employment and the gender digital divide in Latin America: A decomposition analysis. *Telecommunications Policy* 45, 102166.

Gamboa, J., y García, C. (2024). Conectar para incluir: ¿Cuáles son las brechas de alfabetización digital en Perú?. Universidad Nacional Agraria la Molina y Consorcio de Investigación Económica y Social. <https://cies.org.pe/investigacion/conectar-para-incluircuales-son-las-brechas-de-alfabetizacion-digital-en-peru/>

García, J., Medina, P. y Bendezú, L. (2011). Determinantes de la Demanda del Servicio de Internet en la Vivienda y el Rol de la Adquisición de PC's como Limitante del Acceso. OSIPTEL, Gerencia de Políticas Regulatorias. Documento de trabajo N° 11 GPR. <https://repositorio.osiptel.gob.pe/bitstream/handle/20.500.12630/358/Demanda%20de%20Internet%20y%20PCs%20vf%20abr11.pdf?sequence=1&isAllowed=y>

Gavilano, M., y Castillo, C. (2019). Caracterización de la demanda de los servicios públicos de telecomunicaciones 2014-2018. OSIPTEL, Gerencia de Políticas Regulatorias. <https://repositorio.osiptel.gob.pe/handle/20.500.12630/447>

Global System for Mobile Communication Association (2023). Brechas de conectividad en América Latina. <https://www.gsma.com/about-us/regions/latin-america/wp-content/uploads/2023/03/FINAL-Brechas-de-conectividad-en-America-Latina -LONG-report-SPANISH-DIGITAL-30-03-2023.pdf>

Global System for Mobile Communication Association (2024). The State of Mobile Internet Connectivity 2024. <https://www.gsma.com/r/wp-content/uploads/2024/10/The-State-of-Mobile-Internet-Connectivity-Report-2024.pdf>

Hîncu, D., Frătilă, L., & Tantau, A. (2011). Gap indicator for measuring digital divide. *Management Research and Practice*, 3, (2), 74-88. <https://mrp.ase.ro/v03i2/f6.pdf>

Instituto Nacional de Estadística e Información (2024). Perú: Acceso y uso de las tecnologías de la información y comunicación en los hogares y por población, 2011-2023. <https://cdn.www.gob.pe/uploads/document/file/7396285/6306464-peru-acceso-y-uso-de-las-tecnologias-de-informacion-y-comunicacion-en-los-hogares-y-por-la-poblacion-2011-2023.pdf?v=1734614714>

Instituto Nacional de Estadísticas e Informática. (2025). Informe Técnico N° 01 – Marzo 2025. Estadísticas de las Tecnologías de la Información y Comunicación en los Hogares. Trimestre: Octubre-Noviembre-Diciembre 2024. https://www.inei.gob.pe/media/MenuRecursivo/boletines/informe-tecnico_tics_oct-nov-dic24.pdf

International Telecommunication Union (2009-i). Manual para la medición del acceso y del uso de las TIC en los hogares y por las personas. https://www.itu.int/dms_pub/itu-d/opb/ind/d-ind-itcmeas-2009-pdf-s.pdf

International Telecommunication Union (2009-ii) - Measuring the information society. The ICT development index 2009. https://www.itu.int/en/ITU-D/Statistics/Documents/publications/mis2009/MIS2009_w5.pdf

International Telecommunication Union (2012). Impact of Broadband on the Economy. Telecommunication Development Sector, ITU. https://www.itu.int/ITU-D/treg/broadband/ITU-BB-Reports_Impact-of-Broadband-on-the-Economy.pdf

International Telecommunication Union (2018). Economic contribution of broadband, digitization ICT regulation: Econometric modelling for the Americas region.
https://www.itu.int/dms_pub/itu-d/opb/pref/D-PREF-EF.BDT_AM-2019-PDF-E.pdf

International Telecommunication Union (2021-i). The economic impact of broadband and digitization through the COVID-19 pandemic. Econometric modelling.
https://www.itu.int/dms_pub/itu-d/opb/pref/D-PREF-EF.COV_ECO_IMPACT_B-2021-PDF-E.pdf

International Telecommunication Union (2021-ii). The affordability of ICT services 2021. *Policy Brief*. ITU.
https://www.itu.int/en/ITU-D/Statistics/Documents/publications/prices2021/ITU_A4AI_Price_Brief_2021.pdf

Internet Society Foundations (2023). ¿Qué es la equidad digital?.
<https://www.isocfoundation.org/es/2023/09/que-es-la-equidad-digital/>

James, J. (2008). Digital Divide Complacency: Misconceptions and Dangers. *The Information Society*, 24, 54–61.

Jann, B. (2008). The Blinder–Oaxaca decomposition for linear regression models. *The Stata Journal*, 8(4), 453–479.
<https://journals.sagepub.com/doi/pdf/10.1177/1536867X0800800401>

Liao, P., Chang, H., Wang, J., & Sun, L. (2016). What are the determinants of rural-urban digital inequality among schoolchildren in Taiwan? Insights from Blinder-Oaxaca decomposition. *Computers & Education*, 95, 123-133.

Martínez-Domínguez, M., & Mora-Rivera, J. (2020). Internet adoption and usage patterns in rural Mexico. *Technology in Society*, 60, 101226.

<https://www.sciencedirect.com/science/article/pii/S0160791X19302684>

Ministerio de Transportes y Comunicaciones (2011). Plan Nacional para el Desarrollo de la Banda Ancha en el Perú.

<https://www.gob.pe/institucion/mtc/informes-publicaciones/2227526-dgprc-plan-nacional-para-el-desarrollo-de-la-banda-ancha-en-el-peru>

Ministerio de Transportes y Comunicaciones (2023). Innovar para conectar: estrategias y medidas de regulación inteligente para reducir la brecha digital.

<https://cdn.www.gob.pe/uploads/document/file/4341442/Innovar%20para%20conectar.pdf?v=1684376266>

Navas-Sabater, J., Dymond, A., & Juntunen, N. (2002). Telecommunications and Information Services For The Poor. Toward a strategy for universal Access. *World Bank Discussion Paper* N° 432. <https://digitalregulation.org/wp-content/uploads/Telecommunications-and-Information-Services-for-the-Poor.pdf>

Ñopo, H. (2008-i). An extension of the Blinder–Oaxaca decomposition to a continuum of comparison groups. *Economics Letters*, 100, 292–296.

Ñopo, H. (2008-ii). Matching as a Tool to Decompose Wage Gaps. *The Review of Economics and Statistics*, 90(2), 290-299.

https://www.jstor.org/stable/pdf/40043147.pdf?refreqid=fastly-default%3A463110a39e8f243eade5bfa020774f51&ab_segments=&initiator=&acceptTC=1

Oaxaca, R. (1973). Male-female wage differentials in Urban Labor Markets. *International Economic Review*, 14(3), 693-709

<https://www.jstor.org/stable/2525981>

Organisation for Economic Co-operation and Development (2001). Understanding the Digital Divide. *OECD Digital Economy Papers*, No. 49, OECD Publishing, Paris.

<http://dx.doi.org/10.1787/236405667766>

Organisation for Economic Co-operation and Development (2011). *OECD Guide to Measuring the Information Society 2011*, OECD Publishing.

<http://dx.doi.org/10.1787/10.1787/9789264113541-en>

Organisation for Economic Co-operation and Development (2018). Bridging the rural digital divide. *OECD Digital Economy Papers*, No. 265, OECD Publishing, Paris.

<https://dx.doi.org/10.1787/852bd3b9-en>.

Organisation for Economic Co-operation and Development (2021). Bridging digital divides in G20 countries. *OECD Report for the G20. Infrastructure Working Group*.

https://www.oecd.org/content/dam/oecd/en/publications/reports/2021/12/bridging-digital-divides-in-g20-countries_daf5c059/35c1d850-en.pdf

Organismo Supervisor de Inversión Privada en Telecomunicaciones (2023). Informe sobre la brecha digital de género en el uso de internet y el acceso a la telefonía móvil, Gerencia de Políticas Regulatorias. Informe N° 00074-DPRC/2023.

Ospino, C., Roldán, P., & Barraza, N. (2010). Oaxaca-Blinder wage decomposition: Methods, critiques and applications. A literature review. *Revista de Economía del Caribe* N° 5, 237-274. <http://www.scielo.org.co/pdf/ecoca/n5/n5a06.pdf>

Otero, J.V. (2012). Descomposición Oaxaca-Blinder en modelos lineales y no lineales. Instituto L.R.Klein – Centro Gauss Facultad de CC.EE. y EE. Universidad Autónoma de Madrid. N° 20. <https://www.uam.es/uam/media/doc/1606862171313/blinder-oaxaca.pdf>

Park, S.R., Choi, D.Y., & Hong, P. (2015). Club convergence and factors of digital divide across countries. *Technological Forecasting & Social Change*, 96, 92–100.

PRONATEL (2021). Conectividad rural. Una mirada prospectiva del Programa Nacional de Telecomunicaciones. https://cdn.www.gob.pe/uploads/document/file/2372567/Reporte%20Anual%20PRONATEL%20FINAL-DIGITAL3_compressed.pdf.pdf

Rahimi, E. & Hashemi Nazari, S.S. (2021). A detailed explanation and graphical representation of the Blinder-Oaxaca decomposition method with its application in health inequalities. *Emerg Themes Epidemiol*, 18(1), 12. <https://doi.org/10.1186/s12982-021-00100-9>

- Ríos Cahuas, P. (2019). Educación y brecha salarial de género en el Perú. [Tesis de licenciatura, Pontificia Universidad Católica del Perú]. Economía (Lic.).
<http://hdl.handle.net/20.500.12404/14943>
- Robinson, L., Cotton, S., Ono, H., Quan-Haase, A., Mesch, G., Chen, W., Schulz, J., Hale, T., & Stern, M. (2015) Digital inequalities and why they matter. *Information, Communication & Society*, 18(5), 569-582. DOI: <http://dx.doi.org/10.1080/1369118X.2015.1012532>
- Song, M., Orazem, P., & Singh, R. (2006). Broadband access, telecommuting and the urban-rural digital divide. Iowa State University. Department of Economics. Working Paper # 06002.
<https://ideas.repec.org/p/ags/genres/18214.html>
- Stern, P., & Townsend, D. (2007). New models for universal access to telecommunications services in Latin America. REGULATEL.
<https://documents1.worldbank.org/curated/en/647461468263937304/pdf/408290ENGLISH01rsal0Access01PUBLIC1.pdf>
- Telefónica (2024). Inclusión digital. Brecha digital: de la brecha de cobertura a la brecha de uso. Políticas Públicas Digitales. Regulación y Competencia.
<https://www.fundacioncarolina.es/wp-content/uploads/2024/11/politicas-publicas-brecha-digital-posicionamiento-2024-1.pdf>
- Tsetsi, E., & Rains S. (2017). Smartphone Internet access and use: Extending the digital divide and usage gap. *Mobile Media & Communication*, 5(7) 239-255. DOI: 10.1177/2050157917708329.

[https://www.researchgate.net/publication/317569029 Smartphone Internet access and use Extending the digital divide and usage gap](https://www.researchgate.net/publication/317569029_Smartphone_Internet_access_and_use_Extending_the_digital_divide_and_usage_gap)

Valiente, M.R., & Lopez, A.J. (2011). Assessing the regional digital divide across the European Union-27. Fundación de las Cajas de Ahorros. Documento de Trabajo N° 483/2009.

[https://www.researchgate.net/publication/28321344 Assessing the regional digital divide across the European Union-27](https://www.researchgate.net/publication/28321344_Assessing_the_regional_digital_divide_across_the_European_Union-27)

Vargas Montoya, L. (2023). Nuevos enfoques para la medición y descomposición de las brechas digitales en Costa Rica. Universidad de Costa Rica. Facultad de Ciencias Económicas. Instituto de Investigaciones en Ciencias Económicas. Avance del Informe de Investigación del Proyecto PRY01-1395-2022.

https://iice.ucr.ac.cr/conferencia_prensa/infome_brecha.pdf

Vicéns Otero, J. (2012). Descomposición Oaxaca-Blinder en modelos lineales y no lineales. Instituto L.R. Klein – Centro Gauss. Universidad Autónoma de Madrid. Documento de Trabajo N° 20. Marzo 2012.

<https://www.uam.es/uam/media/doc/1606862171313/blinder-oaxaca.pdf>

Vu, K.M. (2011). ICT as a source of economic growth in the information age: Empirical evidence from the 1996–2005 period. *Telecommunications Policy*, 35, 357–372.

<https://doi.org/10.1016/j.telpol.2011.02.008>

Whitacre, B., Strover, S., & Gallardo, R. (2015). How much does broadband infrastructure matter? Decomposing the metro–non-metro adoption gap with the help of the National Broadband Map. *Government Information Quarterly*, 32, 261–269.

W3Techs. (2025). *Usage statistics of content languages of websites*. Datos obtenidos (accedidos el 29 de junio de 2025) en el siguiente enlace.

https://w3techs.com/technologies/overview/content_language

Zhou, Y., Chen, M., Liu, X., & Chen, Y. (2024). A New Framework, Measurement, and Determinants of the Digital Divide in China. *Mathematics*, 12, 2171.

<https://doi.org/10.3390/math12142171>.



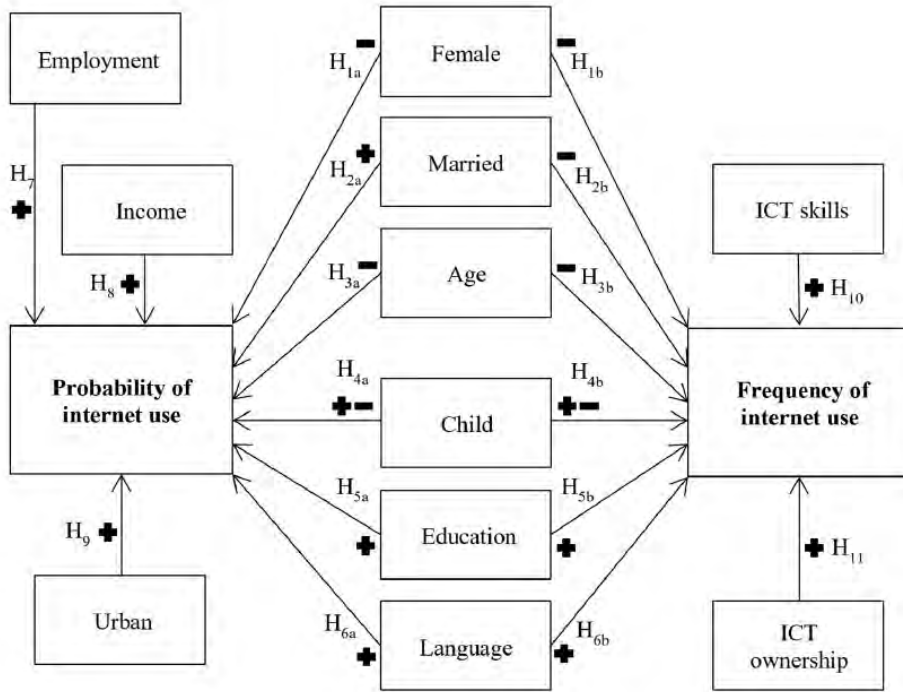
ANEXOS

Anexo 1. Impactos económicos de la banda ancha en la economía. Breve resumen de la literatura

País	Autores	Datos	Efectos
Estados Unidos	Crandall et al. (2007) - Brookings Institution	48 estados de EEUU para el período 2003-2005	Resultados estadísticamente no significativos
	Thompson and Garbacz (2008) - Ohio University	46 estados de EEUU para el período 2001-2005	Un incremento del 10% en la penetración de la banda ancha está asociada con un incremento de 3.6% en la eficiencia
OECD	Czernich et al. 82009) - University of Munich	25 países de la OECD entre 1996 y 2007	Un incremento del 10% en la penetración de la banda ancha, aumenta la tasa de crecimiento del PBI per cápita entre 0.9 y 1.5%
	Koutroumpis (2009) - Imperial College	22 países de la OECD entre 2002 y 2007	Un incremento del 10% en la penetración de la banda ancha, produce un incremento del crecimiento del PBI en 0.25%
Países con ingresos altos	Qjang et al. (2009) - World Bank	66 países con altos ingresos entre 1980 y 2002	Un incremento del 10% en la penetración de la banda ancha, genera un 1.21% adicional en la tasa de crecimiento del PBI
Países con ingresos medios y bajos	Qjang et al. (2009) - World Bank	restantes 120 países (con ingresos medios y bajos) entre 1980 y 2002	Un incremento del 10% en la penetración de la banda ancha, genera un 1.38% adicional en la tasa de crecimiento del PBI

Elaboración propia a partir de ITU (2012), pg 4.

Anexo 2. Efectos esperados de las variables relevantes en el uso del Internet



Elaboración propia a partir de Doyar et al. (2023), Fig 5, pg 9.

Anexo 3. Descomposición Oaxaca-Blinder³⁰

Se tiene la siguiente regresión:

$$y_i = \beta_0 + \sum_{k=1}^K \beta_k x_{ik} + \varepsilon_i$$

Luego se define el índice grupos “g”, donde $g = U, R$ y donde U= urbano y R = rural.

Sobre la base de la ecuación de regresión y los grupos definidos previamente se plantea la regresión en medias siguiente:

$$\bar{Y}^g = \beta_o^g + \sum_{j=1}^K \beta_j^g \bar{x}_j^g$$

A partir de la regresión en medias se plantea la regresión de la diferencia en medias entre los 2 grupos (U y R):

$$\Delta \bar{Y} = \bar{Y}^U - \bar{Y}^R = \beta_o^U + \sum_{j=1}^K \beta_j^U \bar{x}_j^U - \beta_o^R - \sum_{j=1}^K \beta_j^R \bar{x}_j^R$$

Reordenando, se tiene la siguiente ecuación:

$$\Delta \bar{Y} = (\beta_o^U - \beta_o^R) + \sum_{j=1}^K (\beta_j^U \bar{x}_j^U - \beta_j^R \bar{x}_j^R) \dots\dots\dots(1)$$

A partir de la ecuación (1) se tiene que la diferencia en medias entre los 2 grupos (U y R) es la suma de los efectos de diferentes componentes: las diferencias básicas entre ambos grupos que incluye el efecto de variables desconocidas que no están incluidas en el modelo (primer sumando del lado derecho de la igualdad); y la diferencia promedio entre el nivel de cada variable

³⁰ Se ha tomado como referencia a Rahimi et al. (2021).

observable (x_j), así como el efecto diferencial (β_j) de esas variables en los 2 grupos que se comparan (ambos recogidos en el segundo sumando del lado derecho de la igualdad).

Luego cabría preguntarse ¿cuánto (qué tan grande) es la contribución de cada uno de los componentes del modelo a esta diferencia ($\Delta\bar{Y}$)?

Para ello se puede aplicar un enfoque contrafactual a través del cual los coeficientes y variables de nivel de la ecuación de un grupo son reemplazados con los correspondientes valores del otro grupo (llamado grupo de referencia). En este procedimiento, los niveles de las variables explicativas y los coeficientes de regresión en los 2 grupos se suponen alternativamente idénticos para lograr el efecto neto de cada componente.

Entonces, asumiendo que el grupo U (urbano) es el grupo de referencia, se tiene que:

$$\bar{Y}^U = \beta_o^U + \sum_{j=1}^K \beta_j^U \bar{x}_j^U$$

Hacemos que:

$\beta_j^U = \beta_j^R + (\beta_j^U - \beta_j^R)$ y $\bar{x}_j^U = \bar{x}_j^R + (\bar{x}_j^U - \bar{x}_j^R)$, y los reemplazamos en la expresión anterior:

$$\bar{Y}^U = \beta_o^U + \sum_{j=1}^K [\beta_j^R + (\beta_j^U - \beta_j^R)] \sum_{j=1}^K [\bar{x}_j^R + (\bar{x}_j^U - \bar{x}_j^R)]$$

Reemplazando \bar{Y}^U en la ecuación (1) se tiene:

$$\Delta\bar{Y} = (\beta_o^U - \beta_o^R) + \sum_{j=1}^K [\beta_j^R + (\beta_j^U - \beta_j^R)] [\bar{x}_j^R + (\bar{x}_j^U - \bar{x}_j^R)] - \sum_{j=1}^K \beta_j^R \bar{x}_j^R$$

$$\begin{aligned} \Delta\bar{Y} &= (\beta_o^U - \beta_o^R) + \sum_{j=1}^K \beta_j^R (\bar{x}_j^U - \bar{x}_j^R) + \sum_{j=1}^K \beta_j^R \bar{x}_j^R + \sum_{j=1}^K (\beta_j^U - \beta_j^R) \bar{x}_j^R - \sum_{j=1}^K \beta_j^R \bar{x}_j^R \\ &\quad + \sum_{j=1}^K (\beta_j^U - \beta_j^R) (\bar{x}_j^U - \bar{x}_j^R) \end{aligned}$$

Simplificando $\sum_{j=1}^K \beta_j^R \bar{x}_j^R$ se obtiene:

$$\Delta \bar{Y} = \underbrace{(\beta_o^U - \beta_o^R)}_B + \underbrace{\sum_{j=1}^K \beta_j^R (\bar{x}_j^U - \bar{x}_j^R)}_D + \underbrace{\sum_{j=1}^K (\beta_j^U - \beta_j^R) \bar{x}_j^R}_C + \underbrace{\sum_{j=1}^K (\beta_j^U - \beta_j^R) (\bar{x}_j^U - \bar{x}_j^R)}_I \dots (2)$$

Donde:

- El componente B es atribuido a las diferencias básicas entre grupos, e incluye los efectos de las variables desconocidas que no están incluidas en el modelo.
- El componente D hace referencia a los cambios esperados medios en las zonas rurales (grupo R) cuando se le comparan con (cumplen con) el nivel de las covariables de las zonas urbanas (grupo U). Es decir, este componente es la parte de la brecha que se explica por las diferencias de grupo en el nivel de las variables explicativas observables (es el componente explicado). Es conocido como el “efecto (de las) dotaciones” de cada grupo.
- El componente C es una parte de la brecha que representa un cambio esperado medio en el grupo rural cuando se le compara con (cumple con) los coeficientes de regresión del grupo urbano. Por tanto, hace referencia a la parte de la brecha que es causada por el efecto diferencial de las variables observables (este efecto diferencial se traduce en los β de cada regresión) sobre los resultados en los 2 grupos que se comparan. De esta manera esta parte de la brecha no puede ser explicada por el nivel de las variables explicativas que son observables. Es el componente no explicado; y se le conoce como el “efecto (de los) coeficientes”.
- El cuarto componente I hace referencia a la interacción debido al efecto simultáneo de las diferencias en las dotaciones y en los coeficientes.

Nótese que, al igual que el componente C, el componente B recoge las diferencias entre los 2 grupos que no pueden ser explicadas por las covariables observadas x (de hecho, esta diferencia se debe a variables no observables). Por lo tanto, si bien es posible juntar esos dos componentes

(B y C) en una parte “no explicada” (NE), a efectos de hacer consistente la presentación formal de la descomposición Oaxaca-Blinder con los resultados que arroja el paquete econométrico utilizado (Stata), no se procederá a juntar ambas partes.

Hasta este momento se ha supuesto que de los 2 grupos, uno de ellos enfrenta efectos discriminatorios (desventaja) respecto del grupo de referencia. Pero ¿cómo se reexpresaría la brecha $\Delta\bar{Y}$ en la ecuación (3), si ahora existe una condición no discriminatoria, recogida en un vector de coeficientes no discriminatorios β_j^* , tal que ambos grupos U y R deseen alcanzarlo?

Se define β_j^* , tal que: $\beta_j^U \geq \beta_j^* \geq \beta_j^R$, o $\beta_j^U \leq \beta_j^* \leq \beta_j^R$. Entonces:

- Si $\beta_j^U > \beta_j^* > \beta_j^R$, existe discriminación positiva a favor del grupo U (y discriminación negativa en contra de grupo R);
- Si $\beta_j^U < \beta_j^* < \beta_j^R$, existe discriminación positiva a favor del grupo R (y discriminación negativa en contra de grupo U).

Entonces, trabajando nuevamente para el grupo U, se define el resultado medio urbano como $\bar{Y}^U = \beta_0^U + \sum_{j=1}^K \beta_j^U \bar{x}_j^U$ y hacemos que $\beta_j^U = \beta_j^* + (\beta_j^U - \beta_j^*)$ y $\bar{x}_j^U = \bar{x}_j^R + (\bar{x}_j^U - \bar{x}_j^R)$, y los reemplazamos en la expresión anterior:

$$\bar{Y}^U = \beta_0^U + \sum_{j=1}^K [\beta_j^* + (\beta_j^U - \beta_j^*)] \sum_{j=1}^K [\bar{x}_j^R + (\bar{x}_j^U - \bar{x}_j^R)]$$

Asimismo, definimos $\beta_j^R = \beta_j^* + (\beta_j^R - \beta_j^*)$

Trabajamos la brecha $\Delta\bar{Y}$:

$$\Delta\bar{Y} = (\beta_0^U - \beta_0^R) + \sum_{j=1}^K [\beta_j^* + (\beta_j^U - \beta_j^*)] [\bar{x}_j^R + (\bar{x}_j^U - \bar{x}_j^R)] - \sum_{j=1}^K [\beta_j^* + (\beta_j^R - \beta_j^*)] \bar{x}_j^R$$

$$\begin{aligned}\Delta\bar{Y} &= (\beta_0^U - \beta_0^R) + \sum_{j=1}^K \beta_j^* (\bar{x}_j^U - \bar{x}_j^R) + \sum_{j=1}^K \beta_j^* \bar{x}_j^R + \sum_{j=1}^K (\beta_j^U - \beta_j^*) \bar{x}_j^R + \sum_{j=1}^K (\beta_j^U - \beta_j^*) (\bar{x}_j^U - \bar{x}_j^R) \\ &\quad - \sum_{j=1}^K \beta_j^* \bar{x}_j^R - \sum_{j=1}^K (\beta_j^R - \beta_j^*) \bar{x}_j^R\end{aligned}$$

Simplificando el componente $\sum_{j=1}^K \beta_j^* \bar{x}_j^R$ tenemos:

$$\begin{aligned}\Delta\bar{Y} &= (\beta_0^U - \beta_0^R) + \sum_{j=1}^K \beta_j^* (\bar{x}_j^U - \bar{x}_j^R) + \sum_{j=1}^K (\beta_j^U - \beta_j^*) \bar{x}_j^R + \sum_{j=1}^K (\beta_j^U - \beta_j^*) (\bar{x}_j^U - \bar{x}_j^R) \\ &\quad - \sum_{j=1}^K (\beta_j^R - \beta_j^*) \bar{x}_j^R\end{aligned}$$

Luego, trabajamos el último sumando de la parte derecha de la igualdad:

$$- \sum_{j=1}^K (\beta_j^R - \beta_j^*) \bar{x}_j^R = + \sum_{j=1}^K (\beta_j^* - \beta_j^R) \bar{x}_j^R$$

Y reemplazamos en la anterior expresión:

$$\begin{aligned}\Delta\bar{Y} &= (\beta_0^U - \beta_0^R) + \sum_{j=1}^K \beta_j^* (\bar{x}_j^U - \bar{x}_j^R) + \sum_{j=1}^K (\beta_j^U - \beta_j^*) \bar{x}_j^R + \sum_{j=1}^K (\beta_j^U - \beta_j^*) (\bar{x}_j^U - \bar{x}_j^R) \\ &\quad + \sum_{j=1}^K (\beta_j^* - \beta_j^R) \bar{x}_j^R\end{aligned}$$

Expandemos el cuarto sumando de la parte derecha de la igualdad:

$$\Delta\bar{Y} = (\beta_0^U - \beta_0^R) + \sum_{j=1}^K \beta_j^* (\bar{x}_j^U - \bar{x}_j^R) + \sum_{j=1}^K (\beta_j^U - \beta_j^*) \bar{x}_j^R + \sum_{j=1}^K (\beta_j^U - \beta_j^*) \bar{x}_j^U - \sum_{j=1}^K (\beta_j^U - \beta_j^*) \bar{x}_j^R + \sum_{j=1}^K (\beta_j^* - \beta_j^R) \bar{x}_j^R$$

Simplificamos $\sum_{j=1}^K (\beta_j^U - \beta_j^*) \bar{x}_j^R$, y se obtiene:

$$\Delta\bar{Y} = \underbrace{\sum_{j=1}^K \beta_j^* (\bar{x}_j^U - \bar{x}_j^R)}_{\substack{\text{Parte Explicada o} \\ \text{Efecto Dotación}}} + \underbrace{(\beta_0^U - \beta_0^R) + \sum_{j=1}^K (\beta_j^U - \beta_j^*) \bar{x}_j^U + \sum_{j=1}^K (\beta_j^* - \beta_j^R) \bar{x}_j^R}_{\substack{\text{Parte No explicada} \\ \text{(incluye el efecto Discriminación)}}} \dots\dots\dots (3)$$

Donde:

- Parte Explicada (Efecto Dotación): es la parte de la brecha que es explicada por las diferencias en los niveles de las variables o características observadas.
- Parte No Explicada: es la parte de la brecha debido a diferencias en los coeficientes. En primer lugar, se tiene la diferencia entre los interceptos de los grupos $(\beta_0^U - \beta_0^R)$, que recoge el efecto de las variables no incluidas en el modelo y que según el enfoque Oaxaca-Blinder se asocia a efectos discriminación o a restricciones estructurales. En segundo lugar, recoge las diferencias de los coeficientes de los grupos (β_j^U, β_j^R) respecto del coeficiente no discriminatorio β_j^* . Por lo tanto, captura las diferencias en el nivel de las variables no observables, así como sus efectos diferenciales (discriminatorios). Nótese que si todas las covariables no observadas se midieran y por tanto se incluyeran en el modelo, este componente únicamente comprendería la diferencia de los coeficientes de cada grupo respecto del coeficiente no discriminatorio. Este componente determina la porción no explicada de la brecha.

Finalmente (y a efectos de tener una comparación directa con los resultados que arroja la descomposición Oaxaca-Blinder en el Stata), sobre la base de la ecuación (3), es posible que únicamente uno de los 2 grupos experimente discriminación y que el vector no discriminatorio β_j^* , sea igual al vector de coeficientes del otro grupo, esto es:

i) $\beta_j^U = \beta_j^* > \beta_j^R$, o

ii) $\beta_j^U < \beta_j^* = \beta_j^R$.

Sea el caso i) $\beta_j^* = \beta_j^U$, entonces la ecuación (3) se re-expresa de la siguiente manera:

$$\Delta\bar{Y} = \sum_{j=1}^K \beta_j^U (\bar{x}_j^U - \bar{x}_j^R) + (\beta_0^U - \beta_0^R) + \sum_{j=1}^K (\beta_j^U - \beta_j^R) \bar{x}_j^U + \sum_{j=1}^K (\beta_j^U - \beta_j^R) \bar{x}_j^R$$

Finalmente se obtiene:

$$\Delta\bar{Y} = \underbrace{\sum_{j=1}^K \beta_j^U (\bar{x}_j^U - \bar{x}_j^R)}_{\text{Explicado}} + \underbrace{(\beta_0^U - \beta_0^R) + \sum_{j=1}^K (\beta_j^U - \beta_j^R) \bar{x}_j^R}_{\text{No Explicado}} \dots\dots\dots (4)$$

Donde:

- El componente Explicado: es la combinación de los componentes D e I en ecuación 2. Si bien en muchos textos a este componente se le denomina el componente explicado, de hecho una parte del mismo (parte I) contiene la diferencia simultánea de los coeficientes y de los niveles de las covariables en ambos grupos. Por eso, si se desea el componente explicado puro (y crudo), es mejor usar la descomposición bajo la forma de la ecuación (2). Está claro que la ecuación (4) es una forma específica de la ecuación (2) donde los componentes D e I han sido integradas de la siguiente manera:

$$\sum_{j=1}^K \beta_j^R (\bar{x}_j^U - \bar{x}_j^R) + \sum_{j=1}^K (\beta_j^U - \beta_j^R) (\bar{x}_j^U - \bar{x}_j^R) = \sum_{j=1}^K \beta_j^U (\bar{x}_j^U - \bar{x}_j^R)$$

Anexo 4. Detalle de las variables utilizadas

CÓDIGO	DESCRIPCIÓN	FUENTE	VARIABLE PCA	
tp_usint	Tasa de población que usa internet (en %)	ENAHO – INEI		
th_tm	Tasa de hogares con acceso a telefonía móvil (en %)	ENAHO – INEI	infraestructura internet	1
th_int	Tasa de hogares con acceso a internet (en %)	ENAHO – INEI		
th_conex	Tasa de hogares con alguna forma de conexión a internet (en %)	ERESTEL – OSIPTEL	pca_inter	
th_rpag	Tasa de hogares con acceso a red pública de agua (en %)	ENAHO – INEI	infraestructura básica	2
th_rpal	Tasa de hogares con acceso a red pública de alcantarillado (en %)	ENAHO – INEI		
th_gn	Tasa de hogares con acceso a gas natural por red de tuberías (en %)	ENAHO – INEI	pca_servbas	
th_elec	Tasa de hogares con acceso a alumbrado eléctrico (en %)	ENAHO – INEI		
pobee_3a18	Tasa de población en edad escolar (de 3 a 18 años) (en %)	ENAHO – INEI	edad	3
th_mees	Tasa de hogares con al menos un miembro en edad escolar (de 3 a 18 años) (en %)	ENAHO – INEI		
pobam_60a	Tasa de población adulta mayor (de 60 años a más) (en %)	ENAHO – INEI	pca_edad	
th_mam	Tasa de hogares con al menos un miembro adulto mayor (de 60 años a más) (en %)	ENAHO – INEI		
av_edjh	Promedio de edad del jefe/a de hogar (en años)	ENAHO – INEI		
tp_eduap	Tasa de población que no supera la educación primaria (en %)	ENAHO – INEI	educación	4
th_medub	Tasa de hogares con algún miembro que no supera la educación primaria (en %)	ENAHO – INEI		
tp_edsup	Tasa de población que tiene algún nivel de educación superior (en %)	ENAHO – INEI	pca_educ	
th_medus	Tasa de hogares con algún miembro con algún nivel de educación superior (en %)	ENAHO – INEI		
Aedujh	Promedio de años de educación del jefe/a de hogar (en años)	ENAHO – INEI		
tp_empinf	Tasa de población con empleo informal (en %)	ENAHO – INEI	vulnerabilidad económica	5
th_pov	Tasa de hogares pobres (en %)	ENAHO – INEI	pca_vuleco	
av_ipch_corr	Promedio del ingreso per cápita del hogar a precios corrientes (en soles)	ENAHO – INEI	capacidad económica	6
av_ipch_cons2023	Promedio del ingreso per cápita del hogar a precios constantes (en soles del año 2023)	ENAHO – INEI		
av_gpch_corr	Promedio del gasto per cápita del hogar a precios corrientes (en soles)	ENAHO – INEI		
av_gpch_cons2023	Promedio del gasto per cápita del hogar a precios constantes (en soles del año 2023)	ENAHO – INEI	pca_capaeco	
pbipc_corr	PBI per cápita a precios corrientes (en soles)	INEI PUBLICACIONES		
pbipc_cons2007	PBI per cápita a precios constantes (en soles del año 2007)	INEI PUBLICACIONES		
p_ccpp_cobmov	Porcentaje de centros poblados con cobertura de telefonía móvil (en %)	NRIP – OSIPTEL	Calidad	7
reclamos_100h	Reclamos por cada 100 habitantes	NRIP - OSIPTEL	pca_cali	
Nsabe	Tasa de población que no usa internet porque dice que no sabe usar internet o computadora (en %)	ERESTEL – OSIPTEL	razones no uso	8
Nhiu	Tasa de población que no usa internet porque dice que no lo necesita, no le interesa o no le ve utilidad (en %)	ERESTEL – OSIPTEL		
Nxpn	Tasa de población que no usa internet porque dice que es muy caro (en %)	ERESTEL – OSIPTEL	ranouso	

Elaboración propia.

Anexo 5. Detalles de las regresiones

Primer Modelo - Modelo Base (brecha de uso explicada solo por infraestructura básica de internet)

A 5.1.1 Para el grupo 1

Model for group 1						
Linear regression		Number of obs		=	72	
		F(1, 70)		=	583.60	
		Prob > F		=	0.0000	
		R-squared		=	0.8233	
		Root MSE		=	4.4122	
tp_usint	Coefficient	Robust std. err.	t	P> t	[95% conf. interval]	
pca_inter_	.6109858	.0252914	24.16	0.000	.5605436	.661428
_cons	24.83686	2.033698	12.21	0.000	20.78078	28.89294

Elaboración propia.

A 5.1.2 Para el grupo 2

Model for group 2						
Linear regression		Number of obs		=	78	
		F(1, 76)		=	203.26	
		Prob > F		=	0.0000	
		R-squared		=	0.7578	
		Root MSE		=	6.7794	
tp_usint	Coefficient	Robust std. err.	t	P> t	[95% conf. interval]	
pca_inter_	.6594661	.0462561	14.26	0.000	.5673391	.7515932
_cons	17.88165	2.548477	7.02	0.000	12.80591	22.95738

Elaboración propia.

A 5.1.3 Descomposición Oaxaca-Blinder

Blinder-Oaxaca decomposition		Number of obs	=	150		
		Model	=	linear		
Group 1: grupo = 0		N of obs 1	=	72		
Group 2: grupo = 1		N of obs 2	=	78		
<p>explained: $(X1 - X2) * b1$ unexplained: $X2 * (b1 - b2)$</p>						
tp_usint	Coefficient	Robust std. err.	z	P> z	[95% conf. interval]	
overall						
group_1	70.83973	1.229946	57.60	0.000	68.42908	73.25038
group_2	50.26028	1.552107	32.38	0.000	47.21821	53.30236
difference	20.57944	1.980355	10.39	0.000	16.69802	24.46087
explained	16.00453	1.800932	8.89	0.000	12.47477	19.53429
unexplained	4.574914	1.179672	3.88	0.000	2.2628	6.887029
explained						
pca_inter_	16.00453	1.800932	8.89	0.000	12.47477	19.53429
unexplained						
pca_inter_	-2.3803	2.590307	-0.92	0.358	-7.457208	2.696608
_cons	6.955214	3.26047	2.13	0.033	.5648114	13.34562

Elaboración propia.

Segundo Modelo - Modelo Intermedio (se incluyen otros factores de infraestructura, así como la edad y educación)

A 5.2.1 Para el grupo 1

Model for group 1						
Linear regression			Number of obs	=	72	
			F(5, 66)	=	118.37	
			Prob > F	=	0.0000	
			R-squared	=	0.8360	
			Root MSE	=	4.3785	
tp_usint	Coefficient	Robust std. err.	t	P> t	[95% conf. interval]	
pca_inter_	.6049296	.0485834	12.45	0.000	.5079296	.7019296
pca_cali_	-.0415825	.0231121	-1.80	0.077	-.0877273	.0045624
pca_servbas_	.1120673	.0607392	1.85	0.070	-.0092025	.233337
pca_edad_	-.0279094	.0436671	-0.64	0.525	-.1150935	.0592748
pca_educ_	-.0680787	.0404682	-1.68	0.097	-.1488761	.0127187
_cons	27.46071	3.346996	8.20	0.000	20.77821	34.1432

Elaboración propia.

A 5.2.2 Para el grupo 2

Model for group 2						
Linear regression			Number of obs	=	78	
			F(5, 72)	=	98.06	
			Prob > F	=	0.0000	
			R-squared	=	0.8418	
			Root MSE	=	5.6292	
tp_usint	Coefficient	Robust std. err.	t	P> t	[95% conf. interval]	
pca_inter_	.74693	.0428832	17.42	0.000	.6614438	.8324162
pca_cali_	-.0074491	.1603092	-0.05	0.963	-.3270197	.3121215
pca_servbas_	-.3022441	.0611744	-4.94	0.000	-.4241931	-.1802951
pca_edad_	.0630625	.035354	1.78	0.079	-.0074144	.1335394
pca_educ_	.1679574	.0559769	3.00	0.004	.0563695	.2795453
_cons	18.22413	11.96814	1.52	0.132	-5.63393	42.08219

Elaboración propia.

A 5.2.3 Descomposición Oaxaca-Blinder

Blinder-Oaxaca decomposition		Number of obs	=	150		
		Model	=	linear		
Group 1: grupo = 0		N of obs 1	=	72		
Group 2: grupo = 1		N of obs 2	=	78		
<p>explained: $(X1 - X2) * b1$ unexplained: $X2 * (b1 - b2)$</p>						
tp_usint	Coefficient	Robust std. err.	z	P> z	[95% conf. interval]	
overall						
group_1	70.83973	1.236006	57.31	0.000	68.4172	73.26225
group_2	50.26028	1.558128	32.26	0.000	47.20641	53.31416
difference	20.57944	1.988838	10.35	0.000	16.68139	24.47749
explained	15.53048	2.216916	7.01	0.000	11.18541	19.87556
unexplained	5.048962	1.7873	2.82	0.005	1.545919	8.552005
explained						
pca_inter_	15.84589	2.090144	7.58	0.000	11.74928	19.9425
pca_cali_	-.0825045	.0924477	-0.89	0.372	-.2636987	.0986896
pca_servbas_	2.416	1.343568	1.80	0.072	-.2173439	5.049344
pca_edad_	.0720206	.1469261	0.49	0.624	-.2159492	.3599904
pca_educ_	-2.720923	1.631207	-1.67	0.095	-5.918031	.4761849
unexplained						
pca_inter_	-6.971974	3.194907	-2.18	0.029	-13.23388	-.7100718
pca_cali_	-2.432476	11.54239	-0.21	0.833	-25.05514	20.19019
pca_servbas_	16.60859	3.526099	4.71	0.000	9.69756	23.51961
pca_edad_	-5.222644	3.230208	-1.62	0.106	-11.55374	1.108448
pca_educ_	-6.169109	1.842444	-3.35	0.001	-9.780233	-2.557985
_cons	9.236578	12.42734	0.74	0.457	-15.12057	33.59372

Elaboración propia.

Tercer Modelo - Modelo Completo I (se incorporan las percepciones de los consumidores).

A 5.3.1 Para el grupo 1

Model for group 1						
Linear regression		Number of obs		= 72		
		F(10, 61)		= 64.72		
		Prob > F		= 0.0000		
		R-squared		= 0.8962		
		Root MSE		= 3.6223		
tp_usint	Coefficient	Robust std. err.	t	P> t	[95% conf. interval]	
pca_inter_	.5021703	.0553924	9.07	0.000	.3914064	.6129342
pca_cali_	.0000917	.0213383	0.00	0.997	-.0425769	.0427603
pca_servbas_	.117563	.0541575	2.17	0.034	.0092685	.2258576
pca_edad_	-.0453775	.0394353	-1.15	0.254	-.1242332	.0334781
pca_educ_	-.0475458	.0569782	-0.83	0.407	-.1614807	.0663891
pca_capaeco_	.2057208	.0397752	5.17	0.000	.1261853	.2852563
pca_vuleco_	.1821097	.0567279	3.21	0.002	.0686753	.2955442
nsabe	.0504036	.1075703	0.47	0.641	-.1646964	.2655037
nniu	.1169369	.0943877	1.24	0.220	-.071803	.3056769
nvpn	.2232864	.1077991	2.07	0.043	.0077288	.4388441
_cons	7.400919	8.496435	0.87	0.387	-9.588753	24.39059

Elaboración propia.

A 5.3.2 Para el grupo 2

Model for group 2						
Linear regression			Number of obs = 78			
			F(10, 67) = 71.12			
			Prob > F = 0.0000			
			R-squared = 0.8747			
			Root MSE = 5.1928			
tp_usint	Coefficient	Robust std. err.	t	P> t	[95% conf. interval]	
pca_inter_	.6105701	.0585302	10.43	0.000	.4937435	.7273968
pca_cali_	.0878781	.1525958	0.58	0.567	-.2167043	.3924606
pca_servbas_	-.1751128	.0772514	-2.27	0.027	-.3293073	-.0209183
pca_edad_	.0285444	.0434615	0.66	0.514	-.0582051	.1152939
pca_educ_	.2350567	.0676517	3.47	0.001	.1000233	.3700901
pca_capaeco_	.5131769	.1829294	2.81	0.007	.1480482	.8783055
pca_vuleco_	.3804817	.1131502	3.36	0.001	.154633	.6063304
nsabe	-.1737208	.0987022	-1.76	0.083	-.3707312	.0232896
nniu	-.1309741	.1135913	-1.15	0.253	-.3577033	.0957551
nxp	-.0307681	.1998904	-0.15	0.878	-.429751	.3682147
_cons	-11.25412	18.2771	-0.62	0.540	-47.73535	25.22712

Elaboración propia.

A 5.3.3 Descomposición Oaxaca-Blinder

Blinder-Oaxaca decomposition		Number of obs	=	150		
Group 1: grupo = 0		Model	=	linear		
Group 2: grupo = 1		N of obs 1	=	72		
		N of obs 2	=	78		
<p>explained: $(X1 - X2) * b1$ unexplained: $X2 * (b1 - b2)$</p>						
tp_usint	Coefficient	Robust std. err.	z	P> z	[95% conf. interval]	
overall						
group_1	70.83973	1.238801	57.18	0.000	68.41172	73.26773
group_2	50.26028	1.564059	32.13	0.000	47.19478	53.32578
difference	20.57944	1.995221	10.31	0.000	16.66888	24.49001
explained	12.70671	2.385938	5.33	0.000	8.030357	17.38306
unexplained	7.872735	1.902151	4.14	0.000	4.144587	11.60088
explained						
pca_inter_	13.15415	1.999954	6.58	0.000	9.234314	17.07399
pca_cali_	.0001819	.0423382	0.00	0.997	-.0827994	.0831631
pca_servbas_	2.534481	1.209464	2.10	0.036	.163974	4.904987
pca_edad_	.1170975	.1839993	0.64	0.525	-.2435344	.4777294
pca_educ_	-1.900278	2.282061	-0.83	0.405	-6.373035	2.572478
pca_capaeco_	5.591836	1.172202	4.77	0.000	3.294363	7.889309
pca_vuleco_	-7.31228	2.321961	-3.15	0.002	-11.86324	-2.76132
nsabe	-.0755722	.1882558	-0.40	0.688	-.4445468	.2934024
nniu	.0170011	.2308715	0.07	0.941	-.4354986	.4695009
nxpn	.5800913	.3478265	1.67	0.095	-.1016362	1.261819
unexplained						
pca_inter_	-5.322242	3.962838	-1.34	0.179	-13.08926	2.444778
pca_cali_	-6.256003	10.98047	-0.57	0.569	-27.77734	15.26533
pca_servbas_	11.73256	3.814246	3.08	0.002	4.256773	19.20834
pca_edad_	-4.243816	3.372085	-1.26	0.208	-10.85298	2.365349
pca_educ_	-7.386182	2.353354	-3.14	0.002	-11.99867	-2.773693
pca_capaeco_	-5.328414	3.257514	-1.64	0.102	-11.71302	1.056195
pca_vuleco_	-14.74601	9.413928	-1.57	0.117	-33.19697	3.704947
nsabe	9.007051	5.874797	1.53	0.125	-2.507339	20.52144
nniu	10.46066	6.237421	1.68	0.094	-1.764464	22.68578
nxpn	1.300106	1.166498	1.11	0.265	-.9861879	3.586401
_cons	18.65503	20.15544	0.93	0.355	-20.8489	58.15897

Elaboración propia.

Cuarto Modelo - Modelo Completo II (se precisa el rol clave del factor demanda por medio de una verificación a través de una medida sintetizada de barreras de percepción)

A 5.4.1 Para el grupo 1

Model for group 1						
Linear regression			Number of obs	=	72	
			F(8, 63)	=	65.80	
			Prob > F	=	0.0000	
			R-squared	=	0.8850	
			Root MSE	=	3.7526	
tp_usint	Coefficient	Robust std. err.	t	P> t	[95% conf. interval]	
pca_inter_	.4929961	.0560318	8.80	0.000	.3810255	.6049668
pca_cali_	.0112078	.0215871	0.52	0.605	-.0319306	.0543461
pca_servbas_	.1454202	.0538796	2.70	0.009	.0377505	.25309
pca_edad_	-.0592411	.0422256	-1.40	0.166	-.1436222	.0251401
pca_educ_	-.0447117	.0543789	-0.82	0.414	-.1533793	.0639558
pca_capaeco_	.2143418	.0401823	5.33	0.000	.1340439	.2946396
pca_vuleco_	.2039669	.0503266	4.05	0.000	.1033972	.3045366
ranouso	.3009846	.2729107	1.10	0.274	-.2443839	.846353
_cons	4.72248	8.513408	0.55	0.581	-12.29021	21.73517

Elaboración propia.

A 5.4.2 Para el grupo 2

Model for group 2						
Linear regression			Number of obs	=	78	
			F(8, 69)	=	88.47	
			Prob > F	=	0.0000	
			R-squared	=	0.8729	
			Root MSE	=	5.1548	
tp_usint	Coefficient	Robust std. err.	t	P> t	[95% conf. interval]	
pca_inter_	.6163893	.0562373	10.96	0.000	.504199	.7285797
pca_cali_	.0827813	.148925	0.56	0.580	-.2143159	.3798786
pca_servbas_	-.183454	.0745534	-2.46	0.016	-.3321839	-.0347241
pca_edad_	.0234072	.0441457	0.53	0.598	-.0646611	.1114755
pca_educ_	.2345358	.0663582	3.53	0.001	.1021548	.3669169
pca_capaeco_	.5311494	.1676782	3.17	0.002	.1966405	.8656583
pca_vuleco_	.3802792	.1081498	3.52	0.001	.1645262	.5960323
ranouso	-.4915017	.2972042	-1.65	0.103	-1.084408	.1014044
_cons	-9.159977	17.51384	-0.52	0.603	-44.09914	25.77919

Elaboración propia.

A 5.4.3 Descomposición Oaxaca-Blinder

Blinder-Oaxaca decomposition		Number of obs	=	150		
		Model	=	linear		
Group 1: grupo = 0		N of obs 1	=	72		
Group 2: grupo = 1		N of obs 2	=	78		
<p>explained: $(X1 - X2) * b1$ unexplained: $X2 * (b1 - b2)$</p>						
tp_usint	Coefficient	Robust std. err.	z	P> z	[95% conf. interval]	
overall						
group_1	70.83973	1.237335	57.25	0.000	68.41459	73.26486
group_2	50.26028	1.561018	32.20	0.000	47.20074	53.31982
difference	20.57944	1.991928	10.33	0.000	16.67534	24.48355
explained	12.19804	2.346227	5.20	0.000	7.599522	16.79656
unexplained	8.381402	1.883731	4.45	0.000	4.689356	12.07345
explained						
pca_inter_	12.91384	1.995023	6.47	0.000	9.003664	16.82401
pca_cali_	.0222376	.0479858	0.46	0.643	-.0718129	.1162881
pca_servbas_	3.13504	1.225421	2.56	0.011	.7332581	5.536821
pca_edad_	.1528726	.2278723	0.67	0.502	-.2937488	.5994941
pca_educ_	-1.787008	2.177823	-0.82	0.412	-6.055463	2.481447
pca_capaeco_	5.826167	1.189814	4.90	0.000	3.494175	8.158159
pca_vuleco_	-8.189913	2.082861	-3.93	0.000	-12.27225	-4.10758
ranouso	.1248099	.1601694	0.78	0.436	-.1891165	.4387362
unexplained						
pca_inter_	-6.058392	3.905897	-1.55	0.121	-13.71381	1.597025
pca_cali_	-5.100607	10.72396	-0.48	0.634	-26.11918	15.91797
pca_servbas_	13.18365	3.729129	3.54	0.000	5.874693	20.49261
pca_edad_	-4.744793	3.510627	-1.35	0.177	-11.6255	2.135909
pca_educ_	-7.298496	2.284198	-3.20	0.001	-11.77544	-2.821551
pca_capaeco_	-5.490484	3.003401	-1.83	0.068	-11.37704	.3960735
pca_vuleco_	-13.10621	8.871351	-1.48	0.140	-30.49374	4.281319
ranouso	23.11428	11.77073	1.96	0.050	.0440613	46.18449
_cons	13.88246	19.47339	0.71	0.476	-24.28468	52.0496

Elaboración propia.

Quinto Modelo - Modelo Parsimonioso (eliminación de variables estadísticamente no significativas al 10%)

A 5.5.1 Para el grupo 1

Model for group 1						
Linear regression			Number of obs	=	72	
			F(6, 65)	=	74.72	
			Prob > F	=	0.0000	
			R-squared	=	0.8795	
			Root MSE	=	3.7821	
tp_usint	Coefficient	Robust std. err.	t	P> t	[95% conf. interval]	
pca_inter_	.5246816	.0454455	11.55	0.000	.4339208	.6154424
pca_servbas_	.0834372	.0329065	2.54	0.014	.0177183	.1491561
pca_educ_	-.0683187	.0546426	-1.25	0.216	-.1774476	.0408101
pca_capaeco_	.1938461	.0343384	5.65	0.000	.1252677	.2624246
pca_vuleco_	.1816145	.0504284	3.60	0.001	.0809021	.282327
ranouso	.366601	.2717431	1.35	0.182	-.1761074	.9093094
_cons	5.026436	7.432459	0.68	0.501	-9.81721	19.87008

Elaboración propia.

A 5.5.2 Para el grupo 2

Model for group 2						
Linear regression			Number of obs	=	78	
			F(6, 71)	=	128.04	
			Prob > F	=	0.0000	
			R-squared	=	0.8719	
			Root MSE	=	5.1008	
tp_usint	Coefficient	Robust std. err.	t	P> t	[95% conf. interval]	
pca_inter_	.6073529	.0514025	11.82	0.000	.5048592	.7098466
pca_servbas_	-.1523462	.057305	-2.66	0.010	-.266609	-.0380833
pca_educ_	.231694	.0514936	4.50	0.000	.1290187	.3343692
pca_capaeco_	.5095646	.1650807	3.09	0.003	.1804031	.8387261
pca_vuleco_	.3789829	.0951953	3.98	0.000	.189169	.5687969
ranouso	-.5402642	.2944232	-1.83	0.071	-1.127327	.046799
_cons	-.7532526	12.35618	-0.06	0.952	-25.39077	23.88427

Elaboración propia.

A 5.5.3 Descomposición Oaxaca-Blinder

Blinder-Oaxaca decomposition		Number of obs	=	150		
		Model	=	linear		
Group 1: grupo = 0		N of obs 1	=	72		
Group 2: grupo = 1		N of obs 2	=	78		
<p>explained: $(X1 - X2) * b1$ unexplained: $X2 * (b1 - b2)$</p>						
tp_usint	Coefficient	Robust std. err.	z	P> z	[95% conf. interval]	
overall						
group_1	70.83973	1.235212	57.35	0.000	68.41876	73.2607
group_2	50.26028	1.558004	32.26	0.000	47.20665	53.31391
difference	20.57944	1.988247	10.35	0.000	16.68255	24.47634
explained	10.94077	2.239157	4.89	0.000	6.552106	15.32944
unexplained	9.638671	1.756731	5.49	0.000	6.195541	13.0818
explained						
pca_inter_	13.74383	1.86688	7.36	0.000	10.08481	17.40284
pca_servbas_	1.798779	.743943	2.42	0.016	.3406772	3.25688
pca_educ_	-2.730516	2.194232	-1.24	0.213	-7.031133	1.5701
pca_capaeco_	5.269061	1.026325	5.13	0.000	3.257502	7.280621
pca_vuleco_	-7.292396	2.074141	-3.52	0.000	-11.35764	-3.227154
ranouso	.1520191	.1782049	0.85	0.394	-.1972561	.5012944
unexplained						
pca_inter_	-4.059015	3.372936	-1.20	0.229	-10.66985	2.551819
pca_servbas_	9.451897	2.678849	3.53	0.000	4.20145	14.70234
pca_educ_	-7.841219	2.017365	-3.89	0.000	-11.79518	-3.887257
pca_capaeco_	-5.471607	2.937581	-1.86	0.063	-11.22916	.2859452
pca_vuleco_	-14.67142	8.013756	-1.83	0.067	-30.37809	1.035257
ranouso	26.45034	11.68863	2.26	0.024	3.541048	49.35964
_cons	5.779689	14.41931	0.40	0.689	-22.48164	34.04102

Elaboración propia.

Ejemplo de obtención de los resultados de la quinta regresión:

Brecha Total = 20.58

$$\Delta Y = \bar{Y}_{urbano} - \bar{Y}_{rural} = 70.84 - 50.26 = 20.58$$

Parte Explicada (Efecto dotaciones) = 10.94077

$$\left[\begin{aligned} &(\overline{Infra.internet^U} - \overline{Infra.internet^R}) \times \hat{\beta}_{Infra.internet}^U + \\ &(\overline{serv.básicos^U} - \overline{serv.básicos^R}) \times \hat{\beta}_{serv.básicos}^U + \\ &(\overline{educación^U} - \overline{educación^R}) \times \hat{\beta}_{educación}^U + \\ &(\overline{capac.económica^U} - \overline{capac.económica^R}) \times \hat{\beta}_{capac.económica}^U + \\ &(\overline{vulner.económica^U} - \overline{vulner.económica^R}) \times \hat{\beta}_{vulner.económica}^U + \\ &(\overline{barreras.percepción^U} - \overline{barreras.percepción^R}) \times \hat{\beta}_{barreras.percepción}^U \end{aligned} \right.$$

$$\left[\begin{aligned} &(75.293 - 49.09825) \times (0.54247) + \\ &(61.6457 - 40.08721) \times (0.0834) + \\ &(66.10361 - 26.13629) \times (-0.0683) + \\ &(44.51232 - 17.33065) \times (0.19385) + \\ &(34.18203 - 74.33518) \times (0.18161) + \\ &(29.58146 - 29.16679) \times (0.3666) \end{aligned} \right.$$

$$= 13.7438 + 1.7987 - 2.7305 + 5.2691 - 7.2923 + 0.15201 = 10.94077$$

Parte No Explicada (Efecto coeficientes + interceptos) = 9.638676

$$\begin{aligned}
 & \left[\begin{aligned}
 & (\hat{\beta}_{\text{Infra.internet}}^U - \hat{\beta}_{\text{Infra.internet}}^R) \times \overline{\text{Infra.internet}}^R + \\
 & (\hat{\beta}_{\text{serv.básicos}}^U - \hat{\beta}_{\text{serv.básicos}}^R) \times \overline{\text{serv.básicos}}^R + \\
 & (\hat{\beta}_{\text{educación}}^U - \hat{\beta}_{\text{educación}}^R) \times \overline{\text{educación}}^R + \\
 & (\hat{\beta}_{\text{capac.económica}}^U - \hat{\beta}_{\text{capac.económica}}^R) \times \overline{\text{capac.económica}}^R + \\
 & (\hat{\beta}_{\text{vulner.económica}}^U - \hat{\beta}_{\text{vulner.económica}}^R) \times \overline{\text{vulner.económica}}^R + \\
 & (\hat{\beta}_{\text{barreras.percepción}}^U - \hat{\beta}_{\text{barreras.percepción}}^R) \times \overline{\text{barreras.percepción}}^R +
 \end{aligned} \right. \left. \begin{array}{l} \\ \\ \\ \\ \\ \\ \end{array} \right\} \text{Efecto dotaciones} \\
 & \left. \left[(\hat{\beta}_0^U - \hat{\beta}_0^R) \right] \right\} \text{Efecto interceptos}
 \end{aligned}$$

Parte No Explicada (Diferencial en Pendientes) o efecto dotaciones = 3.85898

$$\begin{aligned}
 & \left[\begin{aligned}
 & (49.09825) \times (0.52468 - 0.60735) + \\
 & (40.08721) \times (0.08343 - (-0.152345)) + \\
 & (26.13629) \times (-0.06832 - 0.231694) + \\
 & (17.33065) \times (0.193846 - 0.509565) + \\
 & (74.33518) \times (0.181615 - 0.378983) + \\
 & (29.16679) \times (0.3666 - (-0.54026)) +
 \end{aligned} \right. \\
 & = -4.059016 + 9.451898 - 7.841219 - 5.471607 - 14.67142 + 26.45035 = 3.85898
 \end{aligned}$$

Parte No Explicada (Diferencial en Interceptos) = 5.779688

$$= (5.026436 - (-0.753253)) = 5.779688$$