

# REVISTA IECOS

NÚMERO 18 - AÑO 2017



Instituto de Investigaciones  
Económicas y Sociales  
FIEECS UNI



**UNIVERSIDAD NACIONAL DE INGENIERÍA**  
FACULTAD DE INGENIERÍA ECONÓMICA,  
ESTADÍSTICA Y CIENCIAS SOCIALES

**Revista IECOS**  
**Instituto de Investigaciones**  
**Económicas y Sociales**  
**FIEECS-UNI**

Número 18 - Año 2017

**RECTOR UNI**

Dr. Jorge Alva Hurtado

**DECANO FIEECS**

Dr. Raymundo Arnao Rondán

**DIRECTOR IECOS**

Mg. Carlos García Sandoval

**COMITÉ DE INVESTIGACIÓN**

Dr. Víctor Carranza Elguera

Dra. Edith Aranda Dioses

Dr. Hernán Garrafa Aragón

**EDICIÓN**

Mg. Nilton Zelada Minaya

Hecho del depósito legal en la  
Biblioteca Nacional del Perú N° 2008-05673

**Editado por**

**Instituto de Investigaciones Económicas y Sociales**  
**FIEECS-UNI**

Avenida Túpac Amaru 210, Lima - Perú.

Central telefónica: 481-1070 anexo 5410. Teléfono directo: 382-6032

Correo-e: [iecos@uni.edu.pe](mailto:iecos@uni.edu.pe), página web: [www.fieecs.uni.edu.pe](http://www.fieecs.uni.edu.pe)

**Impreso por**

Xxxxx

# Contenido

In memoriam al maestro Julio Ismodes	5
RAFAEL CAPARÓ Estimación de la demanda oculta de servicios de salud: un apoyo al seguro social peruano	9
RAFAEL CAPARÓ Y ELIZABETH PARI Modelos de econometría espacial para la lucha contra la delincuencia en el Perú: un enfoque de optimización en tiempo real	41
EDUARDO QUIROZ VERA Calidad de la enseñanza universitaria y satisfacción del estudiante de la FIECS UNI (2016)	73
RENÁN JESÚS QUISPE LLANOS Análisis de la pobreza y distribución del ingreso 2009 -2016	98
CARLOS RISCO FRANCO Modelos ARIMA para el Análisis de datos de sismos en el Perú en el 2017	144
JOSÉ P. SAAVEDRA PACHECO Competencias laborales y demanda laboral de los egresados de Ingeniería Económica, UNI 2017	159
WILSON HERNÁNDEZ BREÑA Crimen en Lima: una aproximación con datos distritales	193
JOSÉ CARLOS ORIHUELA, CÉSAR HUAROTO Y CARLOS A. PÉREZ Pasivos ambientales mineros, salud y agricultura: una primera aproximación espacial	239



## In memoriam al maestro Julio Ísmodes Alegría

El 11 de junio de 2017 falleció el profesor Julio Ísmodes, recordado miembro de la Quinta Promoción de la antigua Escuela de Economía de la UNI. Decidido a convertirse en un profesional, superó la estrechez económica compartiendo sus estudios universitarios con la actividad laboral a tiempo parcial en el Jockey Club del Perú.

Una de sus más ansiadas metas personales –y a la que dedicó los últimos años de su vida académica– fue la elaboración de su tesis doctoral sobre la Economía urbana de Lima. Su partida se produjo cuando se encontraba ad portas de realizar su más preciado sueño.

Economía urbana y Proyectos de inversión fueron las especialidades que marcaron su preferencia académica y docente, y definieron incluso las obras que publicaría como autor o coautor: *Economía Urbana. Teoría y Práctica* (1997), *Conceptos básicos de proyectos de inversión* (1991), *Los parques industriales de desarrollo* (1992) y *Desarrollo Empresarial* (1992).

Otro sueño de nuestro maestro fue el impulso de la profesionalización de la carrera de Ingeniería Económica. Julio Ísmodes, Enrique Juscamayta y Alberto Quintanilla – este último congresista de la República–, entre otros, fueron entusiastas impulsores de la Ley N° 25391 de 1992, la cual estableció la profesionalización de los ingenieros economistas y su inscripción alternativa en el Colegio de Ingenieros o en el Colegio de Economistas. Posteriormente, participó activamente en la formación del Capítulo de Ingeniería Económica en el Colegio de Ingenieros del Perú y llegó a ser Presidente de este Capítulo en el periodo 1998-1999.

Su pasión por lograr una formación más sólida del ingeniero economista lo llevó a dirigir la Comisión de Elaboración del Nuevo Plan de Estudios de la Escuela Profesional de Ingeniería Económica en el año 2014. El informe de esta Comisión sirvió de base para que la siguiente Comisión culminara la aprobación del Nuevo Plan de Estudios con posterioridad al deceso del profesor Ísmodes. Esta misma política también la aplicó en la Universidad Tecnológica del Perú (UTP), institución de la que llegó a ser director y organizador de la carrera de Ingeniería Económica y Empresarial.

En su vida profesional fue consultor y asesor en Formulación y evaluación de proyectos, Economía, y Desarrollo urbano y regional de importantes entidades tanto públicas como privadas, entre las que citamos a CEPAL, PETROPERU e IPSS.

Su espíritu docente lo convirtió en un símbolo y referente para muchos maestros y para sus estudiantes que siempre vieron en él un profesor ejemplar en sus clases de Economía urbana y Evaluación social de proyectos.

Apreciado Julio Ísmodes: los docentes y estudiantes de la FIEECS te recordarán por siempre por ser consecuente y tenaz en alcanzar tus ideales.

Leoncio Palacios

revista iecos

## **ARTÍCULOS**





# Estimación de la demanda oculta de servicios de salud: un apoyo al seguro social peruano<sup>1</sup>

Rafael Caparó

## Resumen

La importancia del sector salud se relaciona estrechamente con el bienestar social, pues pensar en este sector salud es pensar en una serie de alternativas de mejora. Muchos de los hospitales que se construyen para el sector salud se saturan de demanda ni bien son terminados de construir, este es un problema que puede estar asociado a una percepción errónea de la demanda de servicios de salud y que se puede reflejar en un malestar social si no es bien calculada, en tal sentido construir un modelo que pueda anidar las expectativas de saturación del hospital (logística) con la demanda oculta de servicio de salud es un gran reto para esta investigación. Tal como se ha hecho en otros países, realizaremos un modelo econométrico para determinar el porcentaje adicional que debe considerarse a la hora de construir de un centro de salud debido a la demanda oculta, sin términos de generalidades este modelo puede aplicarse a otros tipos de servicios sociales. Este es el tercer informe y complementa la investigación con cuadros y resultados.

**Palabras clave:** Demanda oculta, servicios de salud, demanda revelada, incidencia, demanda potencial.

## Introducción

Encontrar una alternativa para detectar el porcentaje de demanda oculta de servicios de salud que puedan terminar en una saturación de los hospitales del seguro social peruano es vital para evitar una disminución del bienestar social. El Seguro Social Peruano (EsSalud) tiene un rol importante en esta investigación, la cual pretende ser una herramienta de análisis para los hacedores de políticas de salud.

La estimación de la demanda oculta es importante en la medida que se puede tener una estimación de lo que en realidad va a ser demandando en términos de servicios de salud, según lo revisado en documentos internacionales, el paciente no revela una enfermedad hasta que no presenta los síntomas y decide atenderse, revelando de esta manera su demanda por este tipo de servicios.

La presente investigación pretende contribuir a una mejora en el cálculo de camillas, médicos, enfermeras, camas, y otros componentes importantes a la hora de invertir en la construcción de hospitales. En tal sentido, la investigación pretende construir un modelo para medir el nivel de demanda potencial o demanda oculta de los servicios de salud. En la investigación también se ha estimado el porcentaje de adicional que se debe considerar para la construcción de algún establecimiento de salud, en función directa de la demanda oculta sirviendo de esta manera como aporte a la toma de decisiones para los hacedores de política pública (salud) relacionadas a la optimización de recursos para la construcción de hospitales.

Se usa un modelo de ingeniería económica orientada al desarrollo social mediante modelos de estimación de demandas potenciales que aporten a la mitigación de la saturación de los hospitales del seguro social en una determinada región del Perú. Se han considerado los siguientes insumos: fuentes de información confiables de instituciones gubernamentales de la región donde se desarrolla el estudio, información sobre hospitales y la demanda de salud de los mismos, datos de los macro-hospitales, datos de enfermedades clasificadas por incidencia y alto costo.

El sector salud es uno de los sectores más importantes a nivel nacional, puesto que influye en el bienestar de los peruanos, la idea de la investigación es mostrar un modelo fácil de entender por los hacedores de política sanitaria, que pueda apoyarlos en la toma de decisiones con una visión prospectiva y no retrospectiva.

La investigación se centra en modelar la capacidad de los peruanos de demandar servicios de salud, como se mostrará más adelante, la demanda de servicios de salud está sujeta a características particulares que tienen que ser estimadas por modelos econométricos especiales, más adelante se mostrarán las referencias bibliográficas.

Los beneficiarios directos de la investigación son los hacedores de política puesto que van a contar con un modelo que pueda apoyarlos en el cálculo de sus costos potenciales, estrechamente ligado a la demanda que pueda presentarse por servicios de salud. Cabe resaltar que esta investigación tiene como beneficiarios a toda persona que demande servicios de salud, en la medida que al revelar su demanda encuentren una logística aceptable capaz de solucionar sus problemas.

Para considerar la coherencia teórica es importante tener una referencia internacional, tanto de economías desarrolladas como de economías subdesarrolladas y países emergentes, para este caso hemos considerado estudios en países como México y Colombia. Puede revisar la bibliografía para profundizar.

El artículo está dividido como se muestra a continuación, en la parte dos hacemos una revisión del estado del arte, en la parte tres consideramos un análisis situacional de

EsSalud, en la cuarta parte describimos la metodología, en la parte cinco, presentamos los resultados y finalmente las conclusiones y recomendaciones de la investigación.

## Estado del arte

Se ha revisado inicialmente, dos documentos de investigación desarrollados en México y Ecuador relacionados con la estimación de demandas ocultas, estos documentos internacionales, sirven como base para el desarrollo del marco teórico y para encaminar a toda la investigación hacia la estimación de la demanda oculta de servicios de salud del Seguro Social del Perú (EsSalud), sin embargo la estimación de esta demanda también puede implementarse para otros sectores, compañías o instituciones financieras, tal como muestra el documento de Robalino y Sánchez (2007) donde se realiza una estimación de la demanda de crédito agrícola, para nuestro informe solo nos concentraremos en la estimación de demandas ocultas de servicios de salud.

La estimación de la demanda oculta se realiza mediante diferentes modelos econométricos, de eficiencia y actuariales. Los modelos econométricos básicos relacionan variables independientes con una variable dependiente para estimar la probabilidad que un asegurado demande los servicios de salud, mientras que el modelo de eficiencia se realiza para encontrar los hospitales que sean los más eficientes, las brechas entre oferta y demanda son estimadas con un modelo actuarial.

Según Santos (2006), para el modelo econométrico se consideran variables socio-económicas correspondiente a cada asegurado para estimar la probabilidad que este demande un servicio de salud, este modelo econométrico es usado en la primera etapa, en la segunda etapa se considera un modelo mejorado de acuerdo a la inclusión de variables obtenidas con el primer modelo, esto mejora notablemente el sesgo de selección y fue propuesto por Heckman en los años 80, la variable que mejora el segundo modelo se conoce como la inversa del ratio de Mills. Estos modelos serán desarrollados matemáticamente en los siguientes apartados. Este estudio estimó una demanda revelada de los servicios de salud que representó el 82 %, y una demanda oculta del 13 %, lo que constituye un gran hallazgo del modelo pues se detectan necesidades sanitarias que por causas socioeconómicas no son reveladas. Este estudio también mostró que solo 34 de 75 hospitales estudiados son eficientes al 100 %, es decir que aproximadamente el 60 % del total de hospitales en México serían eficientes.

Similar a este estudio Robalino y Sánchez (2007) proponen el modelo de Heckman para estudiar la demanda oculta, pero de créditos agrícolas, la propuesta adicional de estos autores es incluir el modelo Tobit como alternativa de estimación, si bien este documento investiga otro tipo de demanda, la importancia de considerar las metodologías elaboradas nos permitirá tener un panorama adicional para la estimación de demandas ocultas de servicios de salud en EsSalud.

Para estos autores estudiar la demanda oculta de créditos es sumamente importante porque ello permite conocer las características de esta demanda oculta, para el

caso de EsSalud, al aplicar este mismo criterio permitiría encontrar las características de los asegurados que forman la demanda oculta de los servicios de salud. Este documento encuentra tanto demanda efectiva como demanda oculta, y una proporción menor que no forma parte de la demanda, como resultados se obtuvieron en uno de los dos escenarios que se plantearon un 20 % de demanda oculta un 74 % de demanda efectiva y un 6 % de la población que no pertenecía a la demanda.

## **Técnicas de estimación de la demanda potencial en México**

En este estudio se estima la demanda potencial para enfermedades de alto costo, también llamadas enfermedades catastróficas (EC de ahora en adelante), las enfermedades catastróficas consideradas son las siguientes: Infarto cerebral, Prematurez, Cardiopatías congénitas, Cáncer cérvico-uterino, Artritis reumatoide, VIH/SIDA, Leucemias y linfomas, Sepsis neonatal, Infección respiratoria del R/N.

En México, la metodología utilizada se sustentó en los procedimientos y técnicas del análisis envolvente y modelos de Heckman, aplicados para:

- a) Estimar la demanda potencial de servicios de salud
- b) Determinar la frontera de producción de máxima eficiencia
- c) Detectar las brechas de demanda-oferta de servicios de salud

Estimación de la demanda potencial: se realiza en dos pasos. El documento desarrollado en México para la estimación de la demanda potencial se basa en el modelo de Heckman, quien desarrolló un procedimiento de cálculo en dos pasos:

Primero se estima la probabilidad de que un individuo decida considerarse enfermo, aquí se hace uso de los modelos probit, este primer paso se conoce como la percepción de la enfermedad.

Segundo, se estima el modelo al añadir una variable –la razón inversa de Mills– para definir la ecuación de comportamiento del consumidor.

## **Experiencia de Argentina**

Trías (2004) investiga los principales determinantes de la demanda por servicios de salud (médicos y hospitalarios) y ve su evolución en el tiempo. Los resultados de dicho estudio sirven para un adecuado diseño de políticas de salud infantil. Para nuestro trabajo se utilizaron datos de 1997 a 2001 (los cuales revelan que los hijos primogénitos tienen mayor probabilidad de acceder a ambos servicios de salud) y se muestra la diferencia entre los factores que afectan el contacto (que es la decisión de llevar o no al niño al médico) y la frecuencia (que representa el número de consultas de los servicios médicos). La cantidad de establecimientos de salud y la tenencia de algún tipo de cobertura influyen sobre el contacto, mientras que el tipo de trabajo de los padres afecta la frecuencia.

Se consideran dos enfoques para medir la demanda de salud. El primero es el enfoque tradicional donde la demanda se basa en la teoría del consumidor, es decir, la demanda de servicios médicos está determinada principalmente por el paciente, pero condicionada por el sistema de salud; en esta línea surge el modelo de Grossman (Grossman, 1972) donde la demanda por los servicios médicos es una maximización intertemporal de la utilidad del paciente. En segundo lugar está el enfoque de agencia, que considera asimetrías en la información entre el médico y el paciente en relación con el tipo y cantidad de tratamiento médico necesario. Aquí el proceso se divide en dos etapas: en la primera la persona elige si busca o no un tratamiento, y en la segunda etapa, una vez que el paciente elige tratarse, el médico determina el número de visitas.

Esta investigación utiliza el enfoque de agencia. En la primera etapa del proceso de decisión se determina el contacto; poseer algún tipo de cobertura de salud puede ser importante para incrementar la probabilidad de acudir a recibir algún tratamiento. En la segunda etapa, donde se determina la frecuencia del tratamiento, se podría captar el comportamiento del médico; aquí podría existir diferencias respecto a que el servicio es público o privado, pues de tratarse de un servicio privado el médico buscaría inducir al paciente a recibir un tratamiento, lo que no ocurre en el sector público.

Para el modelo de dos partes, en la primera etapa se modela la división entre usuarios (positivo) y no usuarios (cero), para ello usamos un modelo de regresión binaria, es decir buscamos en esta primera etapa estimar la probabilidad de acceder al servicio de salud, para lo que empleamos un modelo probit. En la segunda etapa se modela la utilización o frecuencia para aquellos que registran valores positivos en la primera etapa y se usa el modelo binomial negativo con datos truncados. La estructura de los modelos en dos partes se puede aplicar a variables discretas y continuas.

Las variables o regresores utilizados se pueden agrupar en cinco categorías, de acuerdo a sus características. Del niño consideramos la edad, sexo, (enfermo), si cuenta con seguro de salud. Del padre se considera el género, edad del jefe de hogar, años de educación, si trabaja más de 30 horas, al menos uno de los padres es asalariado. Del hogar se considera género del jefe de hogar, tipo de hogar, número de hijos menores a 12 años, posición relativa del niño y nivel de ingreso. De los aspectos regionales se considera la región a la que pertenece, distancia a un centro de salud, número de hospitales a menos de un kilómetro. Y por último, se consideran las características de la oferta de servicio como el número de establecimientos de salud en su localidad.

En la segunda etapa, para estimar la probabilidad de estar internado se consideran las mismas variables explicativas que se utilizaron en la primera etapa, pero en lugar de emplear la variable número de establecimientos de salud se utiliza el número de camas disponibles per cápita y por provincia.

Luego del análisis previo se concluye que los hijos primogénitos tienen mayor probabilidad de acceder a los servicios médicos y hospitalarios. Por otro lado, los niños más pobres tienen menor probabilidad de consulta al médico y una mayor

probabilidad de acceder a los servicios hospitalarios, y permanecen periodos más prolongados. La tenencia de cobertura aumenta la probabilidad de consulta al igual que el número de establecimientos de salud.

## **Experiencia de Colombia**

Zapata y Bardey (2013) buscan identificar los determinantes sociodemográficos de la demanda por planes voluntarios de salud en Colombia. Este trabajo incluye una aplicación de la metodología propuesta por Henríquez (2010), utilizada para estudiar el caso chileno. En este trabajo se identifican nichos de mercado de planes voluntarios que maximizan la probabilidad de adquisición de dichos planes. Se estima el tamaño de mercado potencial de estos planes. Luego se estima el posible ahorro generado por los planes voluntarios de salud al sistema. Con el fin de entender los determinantes de la demanda por planes voluntarios de salud es necesario conocer la naturaleza de la interacción entre el mercado de seguros voluntario, la salud de las personas y la cobertura obligatoria ofrecida por el sector público (Zapata y Bardey, 2013).

## **Metodología usada para estimar la demanda del mercado de seguros voluntarios de salud**

Henríquez (2006) estima la importancia de algunos determinantes sociodemográficos que afectan la decisión de tomar un seguro privado de salud, o no tomarlo, a través de un modelo probit. Con esta metodología se encuentran las determinantes de la demanda y luego se identifican los nichos de mercado que maximizan la probabilidad de adquisición de los planes voluntarios de salud. Para el caso chileno, el autor utilizó como variable explicativa una variable binaria que toma el valor de uno si la persona está inscrita en un plan complementario de salud y cero en otro caso. Las variables explicativas son: edad, género, estado civil, años de educación, ingreso del jefe de familia, si es trabajador independiente, estado de salud (muy buena, buena, mala, muy mala), si pertenece al área urbana y por último, si su trabajo es permanente.

La misma metodología es usada para el caso colombiano pero con diferentes criterios en el modelo: la variable explicada es de tipo binaria y toma el valor de uno si el individuo está en el plan voluntario a la hora en que fue finalizada la encuesta, de lo contrario toma el valor de cero. Las variables explicativas son: edad, género (que toma el valor de 1 si el hombre y 0 si es mujer), el estado civil, nivel educativo (1 si tiene estudios superiores al nivel secundario y 0 de lo contrario), el salario que es una variable continua, independiente (1 si es trabajador independiente y 0 en otro caso), empleo (1 si la persona cuenta con un tipo de trabajo y 0 de lo contrario); también se consideró las variables estado de salud (muy buena, buena, mala y muy mala) y urbano (1 si vive en área urbana y 0 en otro caso).

Luego de estimar el modelo probit con las variables descritas anteriormente, se procede a dividir la probabilidad de participación en deciles y obtener algunas es-

tadísticas descriptivas únicamente del decil más alto de probabilidad, las cuales se comparan con las estadísticas de la media nacional. Mediante este proceso se identifican ciertas características de nichos de mercado que maximizan la probabilidad de adquisición de planes voluntarios de salud.

Luego de desarrollar el modelo propuesto por Henríquez (2010) para el caso Colombiano se obtienen los siguientes resultados. En general, se ve que las personas que tienen mayor probabilidad de compra de planes voluntarios de salud se encuentran entre los 44 y los 49 años de edad, y en promedio tienen un salario que oscila entre los 2500 y los 3300 dólares y este salario per cápita para cada miembro del hogar esta entre los 930.000 y 1 230. Esto significa que los clientes potenciales pertenecen al octavo, noveno y décimo decil de ingreso. Asimismo tienen un nivel educativo promedio que se encuentra entre tecnológico y universitario con título.

Luego de esto se procede a estimar el tamaño potencial de mercado que comprende a la población más propensa a comprar seguro más la población que tiene el mercado, además de identificar las características de los consumidores y la magnitud del mercado potencial se identificó la localización geográfica que maximiza la probabilidad de compra de los clientes potenciales.

## **Experiencia de España**

En la Unión Europea se suele mostrar que los servicios de salud son casi ilimitados, pero se exige establecer medidas eficaces de asignación de recursos para satisfacer las necesidades de los pacientes. Además de la eficiencia, se deben considerar aspectos como equidad, solidaridad y redistribución, que son muy importantes para los sistemas de seguridad social.

Con referencia a España, puntualmente, Clavero y González (2005) realizan una investigación con el propósito de identificar los factores determinantes de la demanda de asistencia sanitaria en este país. En el sistema de salud español existen problemas financieros y de sostenibilidad del sistema debido al aumento continuo del gasto sanitario público, esto exige que se plantean diversas reformas al sistema. Los problemas que genera el aumento del gasto en servicios de salud se deben, generalmente, al uso de tecnologías costosas y al considerable incremento de la demanda, además de otros factores como el envejecimiento de la población; por lo tanto, se necesita conocer la demanda de los servicios de salud, es decir, se necesita conocer los factores que explican la decisión del paciente sobre el consumo de servicios de salud (Pohlmeier y Ulrich, 1995). La fuente utilizada para obtener esta información es el Panel de Hogares de la Unión Europea (PHOGUE).

## **Metodología usada para estimar la demanda de asistencia sanitaria**

Con el fin de aplicar del modelo econométrico que permita estimar la demanda de asistencia sanitaria se utilizó como variable dependiente el número de visitas

efectuadas al médico general y al especialista en el transcurso de los doce meses anteriores a la entrevista. Las variables independientes se agruparon en cinco categorías: indicadores de salud, factores demográficos, características socioeconómicas o hábitos de vida, otros servicios alternativos o complementarios, y oferta de asistencia sanitaria.

Con referencia a los indicadores de salud se consideró el estado de salud (desde muy buena hasta muy mala), si sufre de un problema crónico o enfermedad. Las características demográficas consideran son edad, sexo y si el asegurado vive con familiares o solo. Las características socioeconómicas incluyen tanto variables del individuo como variables del hogar donde vive; en cuanto a las condiciones del hogar se ha elaborado un indicador mediante el método de componentes principales, donde se consideran problemas de habitabilidad como humedad, goteras podredumbre o contaminación que influyen en la salud y el posterior uso de los servicios de salud. En lo que respecta a las condiciones económicas se considera el nivel educativo y la renta que percibe, dichas variables fueron categorizadas. La categoría otros servicios sanitarios considera además de las consultas al médico general y especialista, las consultas al dentista y el número de noches de estancia hospitalaria.

Con las variables descritas se estiman los modelos binomial negativo y hurdle binomial negativo. El objetivo de estimar este modelo es verificar si la utilización de servicios sanitarios responde a un doble proceso de decisión: por un lado, la decisión de iniciar el contacto con el sistema sanitario y por otro, la frecuencia del número de visitas.

En lo que respecta al análisis econométrico, los modelos de elección discreta permiten tener en cuenta las características de la variable analizada, un elevado número de personas que no acuden al médico y una elevada proporción que solo realiza un número reducido de visitas al año. La comparación permite comprobar que la demanda de consultas al médico general corresponde a una única decisión, representada por el modelo binomial negativo, aunque no se ha podido diferenciar a los individuos que utilizan el servicio y a los que no lo hacen. En cambio, la demanda de consultas al médico especialista es un doble proceso de decisión. Desde el punto de vista de la organización del sistema sanitario español es lógico, puesto mientras las consultas de medicina general son de libre acceso para el paciente, las de acceso al especialista tienen que ser aprobadas por otro profesional sanitario que actúa como agente del paciente.

## **Experiencia de Estados Unidos**

De acuerdo con Meyerhoefer y Zuvekas (2006) la respuesta de los consumidores frente a los precios es fundamental para las propuestas de reforma en marcha que buscan resolver rápidamente los crecientes costos de atención de salud, pero las mejores estimaciones disponibles de las elasticidades precio de la demanda son de



hace más de 25 años, por lo que se busca estimaciones más actuales de la demanda para el tratamiento de la salud mental y física a través de un modelo de demanda de atención de salud que incorpora los costos relevantes que influyen en las decisiones de consumo, excluyendo de su bolsillo los pagos (costo compartido) por servicios ambulatorios y por medicamentos recetados, así como las primas de seguros.

Se suele considerar la demanda de atención médica como una demanda derivada de la salud, y por tanto, depende del estado actual de la tecnología médica. En otras palabras, actualmente los consumidores están comprando paquetes de bienes muy diferentes de lo que compraron décadas atrás. Por otra parte, podrían influir en los niveles de demanda relativos y en la sensibilidad al precio, los cambios en la cobertura institucional para el tratamiento de la salud mental y física y sus medicamentos con receta asociados, provocados por la atención médica administrada y el creciente papel del sector público en los mercados de seguros de salud. Lo que se busca es proporcionar estimaciones actuales de la demanda de tratamiento de la salud física y de la salud mental, y en particular la elasticidad precio de la demanda de los diferentes tipos de tratamiento que incluyen medicamentos.

En la primera etapa del procedimiento se utilizó el modelo probit sobre la muestra de los consumidores que realizan gastos de su propio bolsillo para determinar la probabilidad de enfrentar costos positivos fuera de su bolsillo para la respectiva receta. Las estimaciones se utilizaron posteriormente para generar un resultado imparcial estimado de la variable latente o subyacente en el modelo probit para la muestra de no consumidores. En comparación con la mayoría de los productos básicos y muchos otros servicios, los servicios médicos son difíciles de medir en unidades homogéneas.

No solo hay una gran variación en la calidad de la atención médica, algunos procedimientos médicos tienen un uso muy intensivo de recursos, mientras que otros pueden disponerlos fácilmente. Con el fin de limitar la heterogeneidad en el tipo de servicios prestados se centra el enfoque en la atención ambulatoria y se excluye el tratamiento hospitalario. De esta manera se mide el consumo farmacéutico mediante el número de llenados de medicamentos recetados durante un año.

Por último, es necesario desagregar el tratamiento ambulatorio en la atención de salud mental ambulatoria, atención ambulatoria de salud física, productos farmacéuticos de prescripción relacionados con la salud mental y productos farmacéuticos de prescripción relacionados con la salud física. Asimismo, se estiman cuatro tipos de cuidado de la salud del sistema de la demanda empírica: salud mental, visitas, medicamentos para la salud mental, visitas de salud física y medicamentos de salud física. Para facilitar la interpretación, se computan dos medidas de elasticidad precio diferentes.

El principal objetivo de Meyerhoefer y Zuvekas es proporcionar nuevas estimaciones de la capacidad de respuesta del consumo de servicios sanitarios a los cambios en los precios de atención de salud para que puedan utilizarse en el análisis de políticas. Una de las conclusiones más importantes del modelo es que, según las estimacio-

nes, la demanda de visitas de pacientes ambulatorios de salud mental es la menor (elasticidad-precio) de los últimos 25 años. La estimación puntual de la elasticidad precio para las visitas de salud mental durante el periodo 1996-2003 para la población total de Estados Unidos fue 0.06, inferior incluso al de la estimación de 0.12 para la elasticidad precio de las visitas de salud física que realizaron los autores.

Para diferentes subpoblaciones de asegurados no hubo diferencias estadísticamente significativas en las elasticidades entre las visitas de salud física y mental. Una de las razones por las que nuestras estimaciones de elasticidad son menores que los encontrados por los investigadores de RAND puede ser debido a los cambios estructurales e institucionales a los mercados de atención de salud en los últimos veinte o treinta años, como la difusión de la atención administrada. Por lo tanto, nuestras estimaciones de elasticidad deben interpretarse como condicional en el entorno institucional actual.

## **Experiencia de Chile**

Jofré, Vial y Torche (2001) analizan los determinantes de la elección entre seguro de salud público y privado desde un punto de vista dinámico, para ello utilizan variables que cambian en el tiempo y cortes transversales repetidos. Específicamente se requiere determinar si la tendencia de compra de un seguro de salud puede ser explicada por cambios en el comportamiento de los individuos y por el impacto de factores económicos en el ingreso y precio.

El sistema de salud chileno se caracteriza por la interacción entre un sector público y otro privado en la provisión de seguros y prestaciones de salud. Todos los trabajadores dependientes deben aportar parte de su salario imponible, para lo cual cuentan con dos opciones: optar por el Fondo Nacional de Salud (FONASA) o bien las instituciones de salud previsional (ISAPRES). Este sistema asegura que todos los trabajadores dependientes tengan un seguro de salud y protejan su ingreso ante cualquier necesidad de atención. Si el individuo elige afiliarse al sistema público (FONASA), la cotización destinada a la previsión de su salud alcanza el 7% con un tope de 4.2 U.F. de su sueldo y puede optar por dos modalidades: la de libre elección y la institucional. Sin embargo, existen problemas en el sistema público de salud que generan exceso de demanda, lo cual genera listas de espera y afecta la calidad del servicio. Por otro lado, en el sistema privado (ISAPRE) se puede elegir diferentes planes, con diferentes cargos, que dependen del riesgo de la persona y la cobertura deseada.

## **Metodología usada para estimar la demanda por Seguros de Salud en Chile**

Los autores usan datos de cortes transversales repetidos (CTR), puesto que proveen mayor información que datos de una sola serie de cortes transversales, que son generalmente vistos como inferiores a los datos de panel (esto es, datos de los mismos individuos a lo largo de diferentes periodos) para estimaciones de modelos dinámicos.

El análisis de datos de CTR es interesante también porque une datos microeconómicos con datos agregados y relaciona características de individuos o familias con características específicas de un país. El modelo de cortes transversales repetidos se aplica a un modelo la estimación de mínimos cuadrados para tener estimadores consistentes. Collado (1997) extendió el análisis a un modelo dinámico con variable dependiente rezagada y demostró su consistencia mediante el estimador “within groups” cuando el número de periodos de la muestra muy grande.

Para determinar la proporción de personas que eligen ISAPRES se toma como variable dependiente a los trabajadores dependientes o a los jubilados que cotizan en ISAPRE (a los que se les asigna el valor de uno si cotiza y si cotizan en FONASA se asigna cero) y como variables exógenas la edad del que cotiza, el nivel de ingresos, los años de educación (a partir del último año cursado), la ubicación geográfica (que es una variable dicotómica que toma el valor de uno si la persona vive en zona metropolitana y cero en otro caso), el sexo (toma el valor de uno si es hombre y cero si es mujer), el número de dependientes (que son todos los hijos que están en el plan familiar de seguros), el número de intervenciones quirúrgicas en los últimos tres meses y, por último, las enfermedades o los accidentes (que toman el valor de uno si la persona tuvo enfermedad o accidente en los últimos tres meses y cero en otro caso). Las variables anuales incorporadas al modelo son: gasto público en salud, número de médicos del sistema FONASA, número de hospitales públicos y número de camas, número de hospitales privados y número de camas.

Luego de haber utilizado el método de cortes transversales repetidos para estimar la Ecuación de demanda de seguro de salud privado desde un punto de vista dinámico. Las variables más importantes en la compra de un seguro son las variables ingreso, edad y la ubicación geográfica donde viven. Secundarias, aunque importantes, fueron las variables anuales, que caracterizan a los dos sistemas público y privado. Por otra parte, la edad tiene un efecto negativo, mientras mayor sea la edad promedio del grupo, menor en la proporción de éstos que están en ISAPRE. La importancia de CTR recae en la posibilidad de incluir variables que cambian a través del tiempo y ver la dinámica que se produce entre este tipo de efectos y los de corte transversal propiamente.

## **Experiencia de Costa Rica**

Llanos y Morera (2009) buscan conocer los determinantes que afectan la utilización de servicios de salud en Costa Rica. El conocimiento de las características de esta demanda es un elemento clave en el diseño de políticas de salud, ya que los cambios en estas pueden afectar las estimaciones presupuestarias del seguro público, la planificación de la cartera de servicios, así como consideraciones de equidad y sostenibilidad del seguro de salud. Con el propósito de conocer estos determinantes de la demanda de los servicios de salud, se usan modelos como el de Andersen y Newman.

El modelo relaciona tres tipos de factores para explicar los patrones de utilización o consumo de los servicios de salud: predisponentes, de factibilidad (capacidad para obtenerlos) y de necesidad (nivel de la enfermedad). Algunos de los factores que predisponen la utilización de los servicios de salud son las características sociodemográficas, entre las que se citan edad, sexo, educación, ocupación, clase social y actividad laboral. Con respecto a los factores de factibilidad, se consideran las características individuales de nivel de ingreso o tipo de seguro, las características poblacionales de disponibilidad y accesibilidad de servicios, el tipo de seguro médico en individuos con seguro y nivel de ingreso. Los factores de necesidad son las causas más inmediatas del uso de los servicios y habitualmente se los evalúa de acuerdo al estado de salud percibido u otras variables relacionadas con la salud como los días en la cama o la restricción de la actividad, entre otros.

### **Metodología usada para estimar la demanda de utilización de servicios de salud en Costa Rica**

Para estimar la demanda se utilizaron datos de la primera Encuesta Nacional de Salud para Costa Rica (2006), estudio transversal llevado a cabo en una muestra representativa de 7 523 personas. La ventaja de utilizar esta encuesta es que proporciona información detallada a nivel nacional y regional acerca de la valoración del estado de salud de los individuos. Los servicios de atención en salud conocidos como consultas médicas, hospitalización y urgencias se asociaron con las tres categorías de variables propuestas en el modelo de Andersen y Newman para explicar las diferencias en uso.

Las variables utilizadas en el modelo son: grupos de edad, sexo, estado civil, zona de residencia, región de residencia, educación, ocupación y los quintiles de ingreso per cápita, y están categorizadas de acuerdo a su naturaleza. En la elaboración del modelo se realizó un análisis descriptivo de las variables (medidas de tendencia central, dispersión, proporciones y frecuencias), luego se usó el test de Ji-cuadrado para variables cualitativas dicotómicas y el análisis de homogeneidad; luego el test de Kruskal-Wallis para las variables cuantitativas no normales y ANOVA para aquellas variables cuya distribución se acerca a la normal. Todos los datos procesaron en SPSS.

Con referencia a la demanda de consulta médica, entre los factores que predisponen la utilización de este servicio están el sexo, el nivel educativo y el ingreso per cápita; no se observó diferencias en los grupos de mayor edad. En relación con los factores facilitadores, las variables significativas son región de residencia y condición de aseguramiento. Respecto a los factores de necesidad, el haber permanecido en cama más de día y el padecer una enfermedad crónica, son las explicativas. En cuanto a la hospitalización, se considera como factor que predispone a la variable educación.

Con referencia a la demanda de consulta médica, entre los factores que predisponen la utilización de este servicio están el sexo, el nivel educativo y el ingreso per cápita; no se observó diferencias en los grupos de mayor edad. En relación con los

factores facilitadores, las variables significativas son región de residencia y condición de aseguramiento. Respecto a los factores de necesidad, el haber permanecido en cama más de día y el padecer una enfermedad crónica, son las explicativas. En cuanto a la hospitalización, se considera como factor que predispone a la variable educación.

En general, las variables en el modelo de Andersen y Newman sobre la utilización de servicios de salud, se cumplen para Costa Rica. Sin embargo, la significancia de los factores varía entre los servicios de consulta ambulatoria, hospitalización y urgencias.

## **La experiencia del crédito agrícola en Ecuador**

Robalino y Sánchez (2007) estudiaron las necesidades de financiamiento del sector agrícola mediante un análisis estadístico de la demanda potencial de crédito agrícola para la zona de influencia. Con el empleo de los modelos Tobit y el método de dos etapas de Heckman, se planteó como objetivo identificar las variables significativas que inciden en la obtención de un crédito agrícola.

Para el desarrollo de los modelos se consideran variables que brinden las características que los hogares deben poseer para acceder a un crédito, la variable endógena es el monto de la deuda y las variables exógenas son las características socioeconómicas de las familias, incluida la tasa de interés, etc. El análisis de la demanda potencial es un instrumento muy importante para las instituciones financieras debido a la necesidad que tienen de conocer el mercado potencial de crédito. Por demanda potencial se entiende la identificación de clientes que tienen capacidad de pago y pueden afrontar una deuda; para identificarlos se divide a la población en dos grandes grupos: personas que, en efecto, cuentan con un crédito en el sistema financiero (constituyen la demanda efectiva) y personas que poseen capacidad de pago porque generan excedentes, pero que no cuentan con un crédito (constituyen la demanda potencial). También hay un subgrupo que no genera excedentes y, por lo tanto, no forma parte de la demanda.

Para el análisis se plantean dos escenarios. Escenario 1: familias que generan excedentes, sin importar el monto, y podrían acceder a un financiamiento. Escenario 2: familias que generan excedentes mensuales que deben ser superiores a la cuota promedio que posee un prestamista del mercado formal de crédito; en este escenario, las familias que tienen excedentes menores a la cuota promedio no representan la demanda potencial. En el escenario 1 la demanda efectiva es el 74 % de la población (y es constante para ambos escenarios) y la demanda potencial es del 20 %; el 6 % restante no forma parte de la demanda. Sin embargo, en el escenario 2 la demanda potencial se reduce a 14 %, debido a que no solo es necesario que la familia genere excedentes sino que estos sean superiores a la cuota promedio del mercado de crédito.

Por otro lado, luego del análisis estadístico las variables significativas identificadas, que inciden en la demanda de crédito para la producción agrícola son: el propósito del préstamo, los gastos empleados en la actividad productiva, la tasa de interés y el tipo de tenencia de tierra.

Los modelos econométricos han sido efectivos para estimar la demanda oculta de servicios de salud en México, mientras que en Ecuador estos modelos han permitido estimar las demandas ocultas de los créditos agrícolas. La demanda oculta se halla en el orden del 20%, cifra expresada como un porcentaje de demandantes futuros o de un total de población de demandantes. Para el caso del seguro social en el Perú, esto se relaciona con la demanda oculta o demanda potencial de los servicios de salud (en nuestro estudio este análisis se centrará en las enfermedades crónicas de alto costo).

Se recomienda el tratamiento de los datos previos para formar una base sólida para la estimación de los modelos. El desarrollo del trabajo de investigación depende de la calidad de los datos, para efectos de estimación se recomienda considerar tanto variables intrínsecas como variables extrínsecas; las primeras se refieren al paciente o asegurado (o familia), y las segundas a las variables sociales, económicas o a todas aquellas que afectan la posibilidad de presentar alguna de las enfermedades de alto costo. La elaboración de esta base de datos es una recomendación para diferentes estudios.

## **Diagnostico situacional de EsSalud**

El Seguro Social de Salud, que también puede identificarse con el nombre de EsSalud, es un organismo público descentralizado, creado con la finalidad de dar cobertura a los asegurados y sus derechohabientes a través del otorgamiento de prestaciones de salud, que incluye la prevención, promoción, recuperación y rehabilitación; así como de prestaciones económicas y sociales. Principales funciones:

- a) Administra el régimen contributivo de la seguridad social en salud y otros seguros de riesgos humanos.
- b) Inscribe a los asegurados y entidades empleadoras.
- c) Recauda, fiscaliza, determina y cobra las aportaciones y demás recursos establecidos por ley, y puede delegar o conceder tales funciones, en forma total o parcial, en entidades del Estado o privadas, según las normas legales vigentes.
- d) Invierte los fondos que administra y procura su rentabilidad, seguridad y equilibrio financiero dentro del marco legal correspondiente.
- e) Formula y aprueba sus reglamentos internos, así como otras normas que le permitan ofrecer sus servicios de manera ética, eficiente y competitiva.
- f) Determina los periodos de calificación para el otorgamiento de prestaciones del régimen contributivo de la Seguridad Social en Salud, de acuerdo con las modalidades y condiciones de trabajo.
- g) Desarrolla programas de prevención de la salud ocupacional y riesgos profesionales.

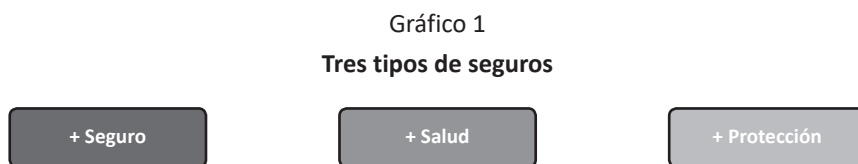
El asegurado y sus derechohabientes:

- El asegurado. Toda persona que se encuentra afiliada, de forma obligatoria o voluntaria, a alguno de los seguros que administra EsSalud para acceder a las prestaciones que estos ofrecen.
- Los derechohabientes. Son aquellas personas que también gozan de beneficios de los seguros que ofrece EsSalud por tener una relación familiar directa con el afiliado titular. Se trata de las siguientes personas:
  - Cónyuge o concubina(o).
  - Hijo(a) menor de edad.
  - Hijo(a) mayor de edad incapacitado(a) en forma total y permanente para el trabajo.

El afiliado titular y/o entidad empleadora debe registrar a los derechohabientes del titular para el acceso a las prestaciones correspondientes.

- Madre gestante de hijo(a) extramatrimonial.

El Seguro Social de Salud administra diversos tipos de seguro, los cuales están agrupados en tres marcas:



*Fuente: Elaboración propia.*

**+Seguro:** Agrupa a los siguientes seguros de carácter obligatorio.

1. Seguro Regular
2. Seguro de Salud Agrario Dependiente

**+Salud:** Agrupa a los siguientes seguros de carácter potestativo.

1. Plan Protección Total
2. Plan Protección Vital
3. Seguro de Salud Agrario Independiente

**+Protección:** Agrupa a los siguientes seguros de carácter complementario a los seguros de EsSalud.

1. Seguro Complementario de Trabajo de Riesgo
2. Vida Seguro de Accidentes

El análisis del entorno demográfico, económico y social en el Perú sirve de base para explicar la dinámica institucional de EsSalud, puesto que permite describir el comportamiento de sus asegurados, así como de su financiamiento.

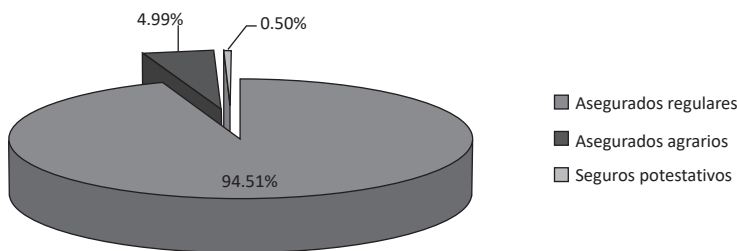
Los requerimientos de un sistema de salud están relacionados directamente con el tamaño, distribución geográfica, estructura por edad y el crecimiento de la población. Se prevé que la población total del Perú crecerá hasta el año 2025 en 1,9 % de promedio anual. Este crecimiento indica importantes desafíos para el fortalecimiento de la protección social en salud en el país.

La realización de estudios de demanda oculta se debe realizar periódicamente para aportar a la gestión de un seguro de salud. Estos estudios permiten contar con información sobre la evolución de los servicios, las demandas potenciales, las fronteras eficientes y las brechas de oferta-demanda; información que servirá como apoyo a la toma de decisiones para la formulación de políticas institucionales en términos de seguridad social.

Se puede identificar tres tipos de asegurados: regulares, agrarios y potestativos. Es importante notar que para setiembre de 2014 los asegurados regulares representaron el 94,5 % del total de asegurados (gráfico 2). Así también para este mismo periodo el total de asegurados –incluidos el titular y los derechohabientes– era de 10 557 997 personas, cantidad que representa un 34,3 % de la población total.

Gráfico 2

**Distribución de asegurados por tipo de seguro  
(Setiembre de 2014)**



Fuente: Gerencia Central de Aseguramiento de EsSalud

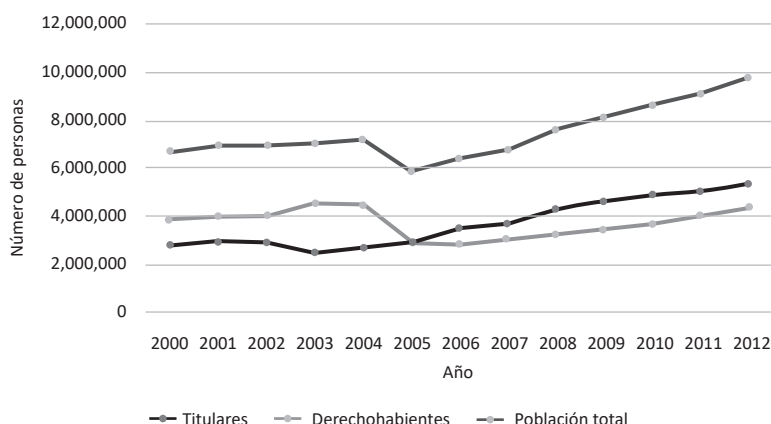
Estos datos proporcionados recientemente por la Gerencia Técnica de EsSalud confirman el análisis realizado y las conclusiones obtenidas en el estudio financiero actuarial del Seguro Social de Salud del Perú (EsSalud) llevado a cabo durante el año 2012 por la Organización Internacional del Trabajo. Esta investigación encontró que para el 2007, la población asegurada representó el 23,19 % de la población nacional, mientras que durante el 2011 ascendió a un 30,5 %; y de mantenerse ese ritmo de crecimiento, será de 44 % para el año 2021.

Además, con respecto a la composición de los asegurados, se puede encontrar que la población total asegurada está en crecimiento, debido principalmente al aumento de titulares, que de derechohabientes. Esto puede explicarse por la disminución de la tasa de natalidad y fecundidad en el país, lo que disminuye el número de derecho-



habientes por titular. Cabe anotar que el estudio encuentra que la población peruana está atravesando un descenso de fecundidad y el aumento de esperanza de vida al nacer, lo que da como resultado un envejecimiento poblacional. Este aumento de la población adulta mayor constituye un desafío para el sistema de seguridad social peruano.

Gráfico 3  
**Evolución de la población asegurada de EsSalud  
 (2000-2012)**



Fuente: Elaboración propia con base en los registros administrativos de EsSalud.

Con respecto a la situación financiera, el estudio encuentra que el estado de ganancias y pérdidas del Fondo Salud que corresponde al seguro regular, es superavitaria, mientras que el Fondo de Seguro Agrario de Salud muestra una situación deficitaria, debido principalmente a que su nivel de aportación es bajo con respecto al seguro regular. Esto concuerda con la prima media general (PMG) obtenida para los distintos tipos de seguro. En el caso del seguro regular la PMG es de 9,08 % y en el caso del seguro de salud agrario es de 9,34 %; y en promedio para regular y agrario es 9,08 %. Esto evidencia que la prima legal (PL) del seguro de salud agrario, que se encuentra en 4 %, no es suficiente para equilibrar sus ingresos con sus gastos.

La Organización Internacional del Trabajo (OIT) realizó un estudio en 2012 para EsSalud. Los resultados que obtuvo son importantes porque sirven de apoyo para la toma de decisiones en favor del cumplimiento de los objetivos propuestos en el Plan Estratégico Institucional 2012-2016 de EsSalud. El diagnóstico que realiza este Plan Estratégico señala que existe una vulnerabilidad de los ingresos a futuro de EsSalud a un mediano plazo, lo cual pone en riesgo la operación de los servicios.

La experiencia muestra que los problemas financieros de EsSalud se deben a la falta de un diseño prospectivo de la sostenibilidad financiera. Sin embargo, también existen

amenazas estructurales a la situación financiera, tales como la disminución del ingreso anual por asegurado y el aumento de los costos de operación hospitalaria. Por otro lado, el aumento de la esperanza de vida, y en consecuencia de los adultos mayores, trae consigo un impacto creciente de carga de enfermedad en la población asegurada, esta situación ha aumentado una fuerte presión de la demanda de servicios, lo que impacta en la situación financiera de la institución. También existe preocupación en torno al acceso a los servicios de salud y la satisfacción por parte de los asegurados.

Finalmente, ambos documentos tienen puntos coincidentes respecto a las medidas que se deben tomar para mejorar la situación de EsSalud. Primero, se carece de estudios actuariales, necesarios para mejorar las primas de determinados seguros. Asimismo, se debe mejorar la gestión de la información, ya que se han encontrado bases de datos que no están integradas. Además, se debe implementar las medidas de fiscalización, recaudación y gestión de deudas necesarias para mejorar la gestión del sistema, para ello se necesita vincular a la SUNAT con los servicios establecidos. Algo también importante, después de haber analizado los distintos seguros que conforman EsSalud, es discutir técnicamente los subsidios cruzados entre cada uno de ellos.

## **El método**

En esta parte describimos el modelo y la metodología para la estimación de la demanda oculta para los seis grupos de diagnósticos CIE10 (Clasificación Internacional de Enfermedades 10) seleccionados previamente. Se considera un modelo revisado en la literatura científica y aplicado tanto al sector de la salud como al sector asegurador.

## **El modelo seleccionado para EsSalud**

Para estimar la demanda oculta de servicios de salud para EsSalud se consideró un modelo secuencial basado en dos etapas –propuesto por Heckman– y orientado a solucionar un problema frecuente en la elaboración de modelos económicos, conocido en el lenguaje técnico como sesgo de selección (Heckman, 1979). Este modelo utiliza los datos para estimar un modelo probabilístico que capture la probabilidad de que se presente una enfermedad, paciente enfermo. A través de un modelo logit se estima la probabilidad que un paciente se considere enfermo. En este paso se captura un ratio conocido como la inversa de Mills, el cual se usa para corregir la primera ecuación ante la presencia de sesgos de selección.

Hasta antes de la publicación del trabajo de Heckman, el error de especificación no constituía un problema grave en el análisis económico, por lo tanto se desconocía que la corrección de dicho error fuera fundamental para obtener conclusiones acertadas y estimadores insesgados, consistentes y eficientes sobre las características de población materia de estudio. En la actualidad, dicho error es uno de los principales problemas que se tiene que tener en cuenta para estimar un modelo y las contribuciones de Heckman han tenido mucha relevancia en cuanto al desa-

rollo de políticas públicas, lo que le llevó a ser premiado con el Nobel en el año 2000 por estos aportes.

El sesgo de selección surge cuando las muestras que disponen los investigadores no son aleatorias, es decir no representan las características reales de la población que es objeto de estudio. Existen diferentes tipos de sesgo de selección que pueden depender del criterio del analista, de los propios agentes económicos etc. El analista al decidir un diseño muestral puede realizar una mala selección de los grupos que se comparan; o sino un problema de autoselección, que se da cuando un individuo se autoselecciona para pertenecer a un determinado grupo. Por lo tanto, para obtener estimaciones no sesgadas se debe tener en cuenta el sesgo de selección.

El método propuesto por Heckman para aislar el sesgo de selección muestral trata este problema en dos etapas. En la primera etapa se estima un modelo tipo logit para calcular la probabilidad del evento que es objeto de estudio y con esta información se estima el ratio de Mills para calcular la magnitud del sesgo; posteriormente al cálculo del modelo de probabilidad, el ratio de Mills estimado se incorpora al modelo de regresión original estimado (MCO) para ser añadido como un regresor más. La significancia del coeficiente del regresor añadido representa el sesgo en que se hubiera incurrido si no se incorporaba este nuevo regresor. De esta manera los coeficientes estimados por MCO son consistentes al aplicar el modelo de dos etapas propuesto por Heckman. Consideramos dos ecuaciones en el modelo, una ecuación de interés que corresponde a la ecuación que se quiere estimar de la cual buscamos sacar conclusiones relevantes, y una ecuación de selección (ecuación auxiliar), que es la que corresponde al modelo de elección discreta, ya sea Logit o Probit, que medirá la probabilidad de estar en la muestra. Este modelo puede incluir variables independientes del modelo de interés y también debe contener al menos una variable continua que determine si se está en la muestra. En ambas ecuaciones a y b, los errores se distribuyen de manera normal.

## **Trabajo de la información**

Se ha revisado los trabajos de Santos (2006), con especial atención al capítulo tres, y Robalino (2007), especialmente los capítulos 5 y 6; de cada documento se analizan párrafos o frases considerados relevantes para el modelo y la investigación. También se ha revisado documentos publicados en la Revista Tecnológica ESPOL.

Se recomienda revisar en el glosario de términos las definiciones de demanda potencial, modelo Tobit, modelo dos etapas de Heckman, modelo Probit, demanda real, variable exógena, variable endógena, eficiencia de factores productivos, proyección financiera, modelo prospectivo, análisis secuencial de la demanda. En la siguiente parte se complementará con mayor detalle las metodologías por considerar.

En cuanto al marco teórico preliminar se realiza una breve descripción de las técnicas de aproximación para la estimación de la demanda, se realiza una revisión del modelo prospectivo para evaluar la sustentabilidad del fondo de protección contra gastos catastróficos en México, donde se resaltan las condiciones del modelo usado.

Se termina esta parte con los resultados de esta aplicación, a continuación se realiza una revisión de modelo diseñado para la estimación de la demanda de crédito agrícola en Ecuador y se pone énfasis, igualmente, en las condiciones del modelo y sus resultados. Este documento termina con el glosario de términos clave.

## **Fundamento teórico: modelos de estimación de demandas de los servicios de salud**

Según la OIT, EsSalud presentó un nivel de cobertura del 30,5 % respecto de la población total del país al año 2011, esto en relación con los más de 10 millones de asegurados frente a los más de 30 millones de peruanos al 2014, bajo la confirmación de la trayectoria de crecimiento reciente; esta cobertura se incrementará al 44 % en el año 2021. Estos pronósticos hacen prioritarios desarrollar un modelo de demanda oculta de los servicios de salud. En esta parte presentaremos toda la base teórica desarrollada internacionalmente para la estimación de demandas de los servicios de salud de EsSalud.

Dentro del fundamento teórico, consideramos los antecedentes y objetivos del informe, el marco conceptual tiene como objetivo presentar los conceptos iniciales en torno a la parte cuantitativa de los métodos y técnicas estadísticas, econométricas y actuariales para desarrollar y determinar los modelos de estimación de la demanda oculta de los servicios de salud que presta EsSalud, en tal sentido se realiza una pequeña presentación de las técnicas de regresión que incorporan los documentos de trabajo, gran porcentaje de este resumen se extrae de documentos internacionales.

A nivel internacional el estudio de la demanda total y la oferta potencial de servicios de salud ha tenido una relevancia importante en los últimos años, debido a su relación con la eficiencia del sector y con la calidad del servicio brindado a los asegurados, así un estudio a profundidad se hace necesario para el Perú.

### **La demanda oculta**

Como toda variable inobservable, la demanda oculta puede ser estimada con métodos estadísticos especializados. La demanda oculta o también llamada demanda potencial puede ser estimada según Heckman (1984) en dos etapas. Para la primera etapa se necesita capturar la probabilidad de que el individuo decida demandar los servicios de EsSalud, esto es una decisión por parte del asegurado. Específicamente, en la primera etapa se estudia la probabilidad de que un individuo decida considerarse como enfermo; en la segunda etapa se construye un modelo para el comportamiento del individuo, y se incluye como una variable explicativa la razón inversa de Mills, esta última variable ayuda a superar el problema de la omisión de información relevante en la recolección de los datos. El primer paso de la estimación corresponde a una técnica conocida como modelación Probit, o

simplemente modelo Probit, que se describe como una probabilidad de éxito y fracaso. De manera simple y probabilística se considera éxito cuando un paciente decide declararse enfermo y se considera fracaso en el caso de que el paciente no decida considerarse enfermo.

## El modelo econométrico

Las dos etapas contienen fundamentos econométricos para la estimación de los parámetros. En esta sección se presentará un modelo econométrico base. Para efectos del modelo inicial, que sirve como punto de partida para la estimación de la demanda potencial, se considera la siguiente relación de probabilidades:

- La probabilidad de que el paciente se considere enfermo será igual a  $Pr(y_i = 1)$ .
- La probabilidad de que el paciente no se considere enfermo será igual a  $Pr(y_i = 0)$ , como caso contrario.

En particular, si suponemos que el indicador  $y_i^*$  depende de las características personales del paciente ( $x'_i beta$ ) y de un componente aleatorio ( $e_i$ ), tendríamos:

$$y_i^* = x'_i beta + e_i$$

Entonces la probabilidad de que el paciente  $i$ -ésimo elija la acción  $y_i = 1$  vendrá dada por:

$$Pr(y_i = 1) = Pr(y_i > c_i) = F(x'_i beta)$$

Donde  $F(\cdot)$  es la función de distribución acumulada de  $y_i$ .

El objetivo es analizar la manera de estimar el vector de parámetros  $beta$  que pondera cada característica personal del paciente, con esto se puede hacer predicciones, aparte del cálculo de la demanda potencial de servicios de salud. Las características personales del paciente sirven como guía para la construcción de las encuestas recolectoras de datos para la estimación de los modelos. Esta investigación usa datos administrativos provenientes de EsSalud que capturan las variables intrínsecas y extrínsecas del modelo.

Un ejemplo de variable intrínseca dentro del modelo son los gastos de una familia en servicios de salud; adicionalmente, se puede definir a un hogar con altos gastos de salud si este destina más del 30 % de sus ingresos al gasto en salud, similar a la definición de enfermedades catastróficas. Las variables adicionales se considerarán dentro del desarrollo de los estudios mencionados.

A continuación presentamos resúmenes de algunos antecedentes internacionales encontrados referente a los temas de investigación.

## **Revisión de las técnicas de aproximación - estimación de la demanda**

En esta parte presentamos una serie de documentos de investigación que proponen diferentes modelos y técnicas estadísticas o de regresión para estimar la demanda, cada documento revisado contiene, además de su resumen, las consideraciones del modelo utilizado y los resultados obtenidos:

### **Modelo prospectivo para evaluar la sustentabilidad del fondo de protección contra gastos catastróficos en México. 2005 - 2010. Hilda Santos**

Con la finalidad de evaluar la sostenibilidad del Fondo de Protección contra Gastos Catastróficos en salud (FPGC) Santos desarrolla metodologías para la estimación de la oferta y la demanda de los servicios de salud además de la eficiencia de los hospitales en la atención de Enfermedades Catastróficas (EC) o Enfermedades de Alto Costo (EAC). En primer lugar para medir la demanda se utiliza un modelo secuencial propuesto por Heckman con dos procesos y/o etapas de cálculo, el primero para estimar la presencia de la enfermedad y el segundo para la demanda del acceso a los servicios médicos. Este último paso nos permite medir no solo la demanda efectiva sino también la demanda oculta, con ello se estima la demanda real de toda la población.

Por otro lado, para estudiar el comportamiento de la oferta total de servicios médicos se analiza tanto la infraestructura instalada, como la eficiencia en brindar los servicios médicos.

Con el propósito de estimar la demanda de los servicios de salud por EC se utiliza el modelo propuesto por Heckman. En la primera etapa se utiliza el modelo probabilístico Probit para estimar la probabilidad de que un persona se reporte como enferma. Se relacionó la variable endógena o dependiente (que es la percepción de la enfermedad), con la utilización de los servicios en los 75 hospitales públicos para las nueve enfermedades catastróficas (EC) seleccionadas. Las variables exógenas o independientes son las características socioeconómicas de la población que utiliza el servicio: sexo, nivel de instrucción e ingreso familiar, las cuales fueron categorizadas. Para la obtención de dichos datos se utilizó una muestra de 1674 encuestas sobre capacidad de respuesta institucional para las EC. Que fueron elaboradas el 2005.

El modelo probabilístico se usa para crear una variable adicional, la inversa del ratio de Mills, que compensa la información omitida en la captación de datos. Una vez obtenida la nueva variable se realiza el segundo modelo para completar el análisis secuencial de la demanda y se calcula la demanda total, es decir no solo la efectiva sino también la demanda oculta o no revelada de la atención médica por las EC.

Para determinar la frontera de producción de eficiencia máxima y estimar la oferta posible de los servicios médicos se usó el análisis envolvente de datos (DEA,

en inglés), con esto se identifica el porcentaje de producción de servicios que pueden y deben aumentar en los hospitales para alcanzar sus niveles potenciales y ello se logra si se utiliza con eficiencia todos los factores. De esta manera se podría aumentar los servicios de atención a los enfermos con EC o EAC para EsSalud. Para calcular los indicadores de eficiencia en cada uno de los 75 hospitales, se utilizó el software Frontier Analyst Professional, versión 3; se asumen como inputs o insumos, el número de médicos en consulta de especialidad seleccionada, el número de enfermeras, la cantidad de camas, el número de consultorios y el de quirófanos. Para la producción obtenida, es decir los outputs, se consideran el número de consultas, las urgencias atendidas, las cirugías en consulta y los egresos por enfermedades catastróficas (EC) para el caso de EsSalud estos egresos serían los gastos anuales en enfermedades de alto costo.

Luego de la estimación de la oferta y la demanda, se realiza el estudio de las brechas entre la oferta y la demanda, para ello se utiliza el software actuarial de la firma Milliman. Para la proyección financiera hasta el 2010 se incluyen tres escenarios: el normativo, el base y el conservador. Asimismo, se consideran modalidades en el escenario normativo, el primero con tarifas al 40 % como se viene operando actualmente, y el otro con el pago completo, lo cual eliminaría el riesgo financiero. Por parte de los hospitales, en cada uno se proyectó cinco diferentes coberturas de las EC, en la primera se consideró las EC con cobertura vigente, y para las cuatro restantes se consideró que para el 2010 de una, dos y tres y finalmente las cuatro EC de mayor prioridad.

## Modelo Tobit

El modelo Tobit es un modelo estadístico que fue propuesto por Tobit en 1958, y se hizo con el propósito de definir la distribución de la variable censurada o truncada, el modelo es usado cuando se tiene datos o información de la variable explicativa, pero no se tienen los datos de la variable explicada, es decir cuando en una muestra la información sobre la variable explicada está disponible solamente para algunas observaciones, a este se le conoce como muestra censurada. Por lo tanto, es aquí donde se utiliza el modelo Tobit, puesto que es capaz de realizar una regresión con datos censurados, la censura es un defecto de los datos de la muestra.

Este método establece un Probit que incorpora tanto la información de la probabilidad del evento de interés, como los valores de la variable subyacente o latente. En términos estadísticos, se puede expresar el modelo Tobit de manera que cumpla con los supuestos del modelo lineal clásico:

$$y_i^* = X_i' \beta + u_i$$
$$y_i = y_i^*, \text{ si } y_i^* > a, \text{ o } y_i = \max(a, y_i^*)$$

Donde:  $u_i \sim N(0, \sigma^2)$

## Interpretación de los parámetros e inferencia

El problema del truncamiento radica en que la variable explicada solo se observa si los valores de  $y$  sobrepasan cierto umbral mínimo  $a$ ; por tanto, para estimar el vector  $\beta$  y  $\sigma$  se necesita conocer la distribución de  $y$  cuando  $y_i^* > a$ , es decir la distribución de probabilidad truncada, y los valores de  $\beta$  y  $\sigma$  se obtienen de la maximización de la función de verosimilitud del modelo censurado Tobit. Amemiya (1973) demostró que el estimador de máxima verosimilitud propuesto por Tobin para este modelo es consistente.

La variable latente cumple con las suposiciones del modelo lineal clásico (MCO), en el caso de la censura por izquierda como se ve en el ejemplo, la variable observada es  $y^*$  cuando  $y_i^* > a$ , y por otro lado es  $a$  cuando  $y_i^* < a$ . La distribución de la variable  $y$  es mixta dado que es discreta cuando  $y_i^* < a$ , y es continua para los demás casos.

Los efectos parciales en el modelo Tobit sobre la variable latente son representados no solo por el valor de  $\beta$ , como suele ocurrir en el modelo lineal general. Para el modelo Tobit en efecto parcial será:

$$d E(Y_i/X_i) / d X_j = \beta_j \Phi(X_i' \beta / \sigma)$$

Donde  $\Phi$  representa la función de distribución acumulada del modelo Tobit, cabe mencionar que la expresión anterior es para el caso de que esté censurado en cero es decir para  $a = 0$ .

Los modelos de regresión censurada son extensiones del modelo Tobit, es decir depende de donde se coloque la censura. Existen varias clasificaciones de las formas de censura del modelo Tobit, Amemiya (1985) y Wooldridge (2002) proponen las siguientes:

- Censura superior.
- Censura inferior y superior.
- Censuras que dependen de características del individuo.
- Modelos censurados en donde la censura o el truncamiento se da respecto de otra variable.

Schnedler (2005) proporciona una fórmula general para obtener estimadores consistentes de probabilidad para estas y otras variaciones del modelo Tobit.

## El modelo de Heckman

El modelo de Heckman está orientado a subsanar un problema que frecuentemente suele ocurrir en la elaboración de modelos económicos, que en el lenguaje técnico se conoce como sesgo de selección. Este problema fue desarrollado por James Heckman en su trabajo Sample selection bias as a specification error (1979).



Hasta antes de la publicación del trabajo de Heckman el error de especificación no constituía un problema grave en el análisis económico, por lo tanto se desconocía que la corrección de dicho error fuera fundamental para obtener conclusiones acertadas y estimadores insesgados, consistentes y eficientes acerca de las características de población que es materia de estudio. En la actualidad dicho error es uno de los principales problemas a tener en cuenta a la hora de estimar un modelo. Las contribuciones de Heckman han tenido mucha relevancia en el desarrollo de políticas públicas, lo que le valió que fuera premiado con el Nobel en el año 2000.

### Sesgo de selección muestral

El sesgo de selección surge cuando las muestras que disponen los investigadores no son aleatorias, es decir no representan las características reales de la población que es objeto de estudio. Existen diferentes tipos de sesgo de selección que pueden depender del criterio del analista, de los propios agentes económicos etc.

El analista al decidir un diseño muestral puede realizar una mala selección de los grupos que se comparan o sino un problema de autoselección, que se da cuando un individuo se autoselecciona para pertenecer a un determinado grupo. Por lo tanto para obtener estimaciones no sesgadas se debe tener en cuenta el sesgo de selección.

El método propuesto por Heckman para aislar el sesgo de selección muestral, consiste en tratar este problema en dos etapas. En la primera etapa se estima un modelo tipo Logit para calcular la probabilidad del evento que es objeto de estudio, con esta información se estima el ratio de Mills para calcular la magnitud del sesgo, posterior al cálculo del modelo de probabilidad, el ratio de Mills estimado se incorpora al modelo de regresión original estimado (MCO) para ser añadido como un regresor más.

La significancia del coeficiente del regresor añadido representa el sesgo en que se hubiera incurrido si no se incorporaba este nuevo regresor. De esta manera los coeficientes estimados por MCO son consistentes al aplicar el modelo de dos etapas propuesto por Heckman. El modelo se puede simplificar de la siguiente manera:

$$y_{2i} = Z_i\delta + v_{2i} \quad (a)$$

$$y_{1i} = X_i\beta + u_{1i}, \quad \text{si } y_{2i} > 0 \quad (b) \quad \text{Se observa si } y_{2i} > 0$$

Consideramos dos ecuaciones en el modelo: una ecuación de interés que corresponde a la ecuación que se quiere estimar y de la cual buscamos sacar conclusiones relevantes; y otra ecuación de selección (ecuación auxiliar) correspondiente al modelo de elección discreta, ya sea Logit o Probit, que medirá la probabilidad de estar en la muestra. Este modelo puede incluir variables independientes del modelo de interés y debe contener al menos una variable continua que determine si se está en la muestra. En ambas ecuaciones a y b los errores se distribuyen de manera normal.

## Revisión de documentos referenciales básicos: estimación de demandas ocultas

La revisión inicial de dos documentos de investigación desarrollados en México y Ecuador relacionados con la estimación de demandas ocultas sirven de base para el desarrollo del marco teórico y para encaminar la investigación hacia la estimación de la demanda oculta de servicios de salud del Seguro Social del Perú (EsSalud). Sin embargo, la estimación de esta demanda también puede implementarse para otros sectores, compañías o instituciones financieras, tal como muestra el documento de Robalino y Sánchez (2007), en el que se realiza una estimación de la demanda de crédito agrícola. Nuestro trabajo se concentra en la estimación de las demandas ocultas de servicios de salud.

La estimación de la demanda oculta se realiza mediante diferentes modelos econométricos, de eficiencia y actuariales. Los modelos econométricos básicos relacionan las variables independientes con una variable dependiente para estimar la probabilidad que un asegurado demande los servicios de salud. Por su parte, el modelo de eficiencia se realiza para hallar los hospitales más eficientes. Las brechas entre oferta y demanda se estiman con un modelo actuarial.

Según Santos (2006), para el modelo econométrico se consideran variables socioeconómicas correspondientes a cada asegurado a fin de estimar la probabilidad que este demande un servicio de salud. Este modelo econométrico se usa en la primera etapa, en la segunda etapa se considera un modelo mejorado de acuerdo a la inclusión de variables obtenidas con el primer modelo, esto disminuye notablemente el sesgo de selección propuesto por Heckman; la variable que mejora el segundo modelo se conoce como la inversa del ratio de Mills. Estos modelos desarrollarán dos matemáticamente. Este estudio estimó una demanda revelada de los servicios de salud que representó el 82 %, y una demanda oculta del 13 %, lo que constituye un gran hallazgo del modelo, pues ha detectado necesidades sanitarias que por causas socioeconómicas no han sido reveladas. Este estudio también mostró que solo 34 de 75 hospitales estudiados son eficientes al 100 %, es decir que aproximadamente el 60 % del total de hospitales en México serían eficientes.

Robalino y Sánchez (2007) proponen el modelo de Heckman para estudiar la demanda oculta de créditos agrícolas. La propuesta adicional de estos autores es incluir el modelo Tobit como alternativa de estimación y si bien este documento investiga otro tipo de demanda, la importancia de considerar las metodologías elaboradas nos permitirá tener un panorama adicional para la estimación de demandas ocultas de servicios de salud en EsSalud, pues para el caso de EsSalud, aplicar este mismo criterio permitiría encontrar las características de los asegurados que forman la demanda oculta de los servicios de salud. El trabajo de Robalino y Sánchez encuentra tanto demanda efectiva como demanda oculta, y una proporción menor que no forma parte de la demanda, es decir, obtuvieron en uno de los dos escenarios planteados:

20% de demanda oculta, 74% de demanda efectiva y 6% de la población que no pertenecía a la demanda.

Con respecto a la estimación de la demanda, al aplicar el análisis secuencial se obtiene que la demanda revelada de los servicios de salud para enfermedades catastróficas (EC) representó el 81,5 % de la demanda total, esto implica que el 12,5 % de la demanda total por los servicios de salud en EC permaneció oculta, lo que constituye un gran hallazgo del modelo, pues se detectan necesidades sanitarias que por causas socioeconómicas no son reveladas. Con todos estos datos se construyó un modelo prospectivo, cuyas conclusiones se sustentan en un conocimiento más preciso de la situación acerca de las EC. Finalmente, se proyectó un escenario para el mercado de EC priorizadas y se incluyó el efecto del crecimiento poblacional.

Por el lado de la oferta de los servicios de salud para las EC, estudiada por Santos (2006), se detectó que solo 34 de los 75 hospitales seleccionados son eficientes al 100% en términos de productividad; para lograr una mejora en los hospitales que no son eficientes se busca replicar las prácticas operativas de los más eficientes. En relación con la proyección financiera se busca analizar la sustentabilidad del Fondo, para ello se proyecta un saldo del fondo de protección contra gastos catastróficos en salud (FPGC) para las cinco modalidades de cobertura, en la modalidad 1 del escenario normativo solo se presenta un escenario con déficit, mientras que en la modalidad 2 se observa que el pago integral de los costos de atención hace insostenible el aspecto financiero del Fondo.

## Resultados

Existe una demanda oculta que satura los servicios de salud de un hospital en el muy corto plazo y genera pérdida de bienestar social. Esta demanda oculta puede ser estimada y minimizada para ser considerada dentro de las expectativas de los hacedores de política, con el fin de mejorar la calidad de los servicios. La demanda oculta se ha estimado por enfermedad crónica de alto costo.

Según los resultados del modelo, la demanda de servicios de salud satura en el muy corto plazo los servicios prestados por los macro hospitales. Es posible encontrar una mejor asignación de los recursos (camillas, salas, centros quirúrgicos, enfermeras, etc.) en función de una demanda de servicios de salud por revelarse, si se proporciona recursos adicionales a los recursos de partida, aproximadamente entre un 20 % y 30 %.

El modelo econométrico basado en datos históricos recoge patrones de demanda de servicios de salud, los cuales convergen a una demanda potencial o una demanda oculta, que solo se presenta cuando el paciente decide revelar la enfermedad. Si no se estima bien esta demanda se puede llegar a tener un mal cálculo de lo que realmente se requiere para satisfacer las necesidades de los pacientes, lo que termina, de esta manera, en pérdidas de bienestar, que provocan una disminución del bienestar social en más del 30 %.

Si se aplican las recomendaciones de aumentar el presupuesto de manera gradual, considerando las enfermedades crónicas de alto costo, la mejora de las redes de hospitales nacionales, así como el monitoreo de los servicios prestados, se produce una mejora del 25% en promedio.

## Conclusiones y recomendaciones

Se trabajó con un conjunto de enfermedades clasificadas de acuerdo a su diagnóstico CIE10, se agruparon de acuerdo a enfermedades crónicas y se estimó la demanda oculta de seis enfermedades previamente seleccionadas como de alto costo. La demanda oculta de los servicios de salud para EsSalud ha sido estimada entre 22% y 27% (a un 95 % de confianza) para el conjunto de enfermedades seleccionadas.

La tabla siguiente muestra que la demanda oculta para el seguro social del Perú tiene una diferencia con la demanda programada de 24,13 % a 25,43 %. De acuerdo con los estudios revisados, este intervalo para la demanda oculta es próximo, al intervalo de la demanda oculta de México, donde el intervalo para las EC es de 18,49 % y 18.52 %.

Tabla 1  
Proyección de mercado para las enfermedades crónicas seleccionadas

	Efectiva	Oculta	Probable	Diferencia (%)
Insuficiencia renal crónica	42591	13549	56140	24.13%
Tumor maligno	100101	32159	132260	24.31%
Hipertensión arterial	410963	139013	549976	25.28%
Diabetes mellitus insumo dependiente	462393	157710	620103	25.43%
Transtorno mental y del comportamiento	298515	99141	397656	24.93%
Músculo-esqueléticas	327181	109174	436355	25.02%

Fuente: Elaboración propia.

En promedio un 24,85 % de la demanda probable sería la demanda oculta para estas seis enfermedades previamente seleccionadas.

Si bien se realiza una estimación de la demanda oculta de manera agrupada para un conjunto de diagnósticos CIE10, se recomienda realizar un análisis por cada diagnóstico, de esta manera se puede estimar demandas ocultas individualizadas y converger hacia mejores políticas preventivas, de monitoreo o de mitigación. En este caso la variable dependiente del modelo de probabilidad tomaría el valor de 1 si el paciente es considerado dentro de un diagnóstico en particular y 0 en otros casos.

## Referencias bibliográficas

- Aracena-Genao, B. et al. (2011). El Fondo de Protección contra Gastos Catastróficos: tendencia, evolución y operación. *Salud pública de México*, 53(4).
- Chang, H-Y. et al. (2010). Comparison of alternative risk adjustment measures for predictive modeling: high risk patient case finding using Taiwan's National Health Insurance claims. *Health Services Research*.
- Clavero y González. (2005). La demanda de asistencia sanitaria en España desde la perspectiva de la decisión del paciente.
- Cohen, J. y Krauss, N. (2014). Spending And Service Use Among People With The Fifteen Most Costly Medical Conditions, 1997. *Health Affairs*. Ellis, R. (1999). Risk adjustment in competitive health plan markets. *Handbook of Health Economics*.
- Crooks, P. (2005). Managing High-Risk, High-Cost Patients: The Southern California Kaiser Permanente Experience in the Medicare ESRD Demonstration Project. *The Permanente Journal*, 9(2).
- Freund, T. et al. (2011). Identification of Patients Likely to Benefit From Care Management Programs. *The American Journal of Managed Care* 17( 5).
- Gunn, J. et al. (2006). *A systematic review of complex system interventions designed to increase recovery from depression in primary care*.
- Heckman, J. (1979). *Sample selection bias as a specification error*.
- Henríquez, R. (2006). Private health utilization and utilization of health services in Chile.
- Jofré, Vial y Torche. (2001). *La demanda por seguros de salud en Chile: una aplicación de cortes transversales repetidos*.
- Kronick, R. et al. (2000). Improving Health-Based Payment for Medicaid Beneficiaries: CDPS. *Health Care Financing Review*.
- Llanos y Morera. (2009). *Factores relacionados con la utilización de servicios de salud en Costa Rica*. Revista Salud Pública.
- Meyerhoefer y Zuvekas, S. (2006). *New Estimates of the Demand for Physical and Mental Health Treatment*.
- Organización Mundial de la Salud (2010). *Informe sobre la salud en el mundo*.
- Panattoni, L. et al. (2011). Predictive risk modelling in health options for New Zealand. *Australian Health Review*, 35, 45-51.

Robalino y Sánchez (2007). *Estimación de la demanda de crédito agrícola para el cantón La Maná, provincia de Cotopaxi, Ecuador.*

Santos, H. (2010). *Modelo prospectivo para evaluar la sustentabilidad del fondo de protección contra gastos catastróficos en México. 2005 - 2010.*

Shi, W, et al. (2011). Effect of household and village characteristics on financial catastrophe and impoverishment due to health care spending in Western and Central Rural China: A multilevel analysis. *Health Research Policy and Systems, 9(16).*

Trías, J. (2004). *Determinantes de la utilización de los servicios de salud: el caso de los niños en la Argentina.*

Xu, K. et al.(2003). Household Catastrophic Health Expenditure: A Multicountry Analysis. *Lancet, 362(9378) 111-117.*

Zapata y Bardey. (2013). *Mercado de seguros voluntarios de salud en Colombia.*

## **Anexo. Glosario de términos**

1. Demanda potencial. También considerada como demanda oculta.
2. Modelo Tobit. Extensión del modelo Probit, pero con un truncamiento.
3. Modelo dos etapas de Heckman. Modelo que en la primera etapa calcula una probabilidad y luego se usa los pronósticos para construir una variable que corrija el sesgo de selección.
4. Modelo probit. Modelo basado en la función probabilística.
5. Demanda efectiva. Es la demanda manifiesta.
6. Demanda oculta. Es la que no se manifiesta, pero que se podría manifestar en un futuro.
7. Variable exógena. Es la que se encuentra fuera del modelo.
8. Variable endógena. Es aquella que el modelo desea explicar.
9. Eficiencia de factores productivos. Situación en que todos los recursos se están utilizando de manera eficiente.
10. Proyección financiera. Flujos de ingresos y egresos futuros.
11. Fondo. Cantidad de dinero administrado con un propósito especial.

12. Población asegurada. Asegurados del Seguro Social de Salud, afiliados regulares, potestativos y sus derechohabientes. Son afiliados regulares los trabajadores activos que laboran bajo relación de dependencia, en calidad de socios de cooperativas de trabajadores; y los pensionistas que reciben pensión de jubilación, de incapacidad o de sobrevivencia.
13. Población asegurada activa. Aquella que tiene como derecho de cobertura al menos los servicios de emergencia. Incluye también a trabajadores eventuales y beneficiarios de prestaciones por desempleo (latencia) por un tiempo no menor de dos meses y hasta por doce meses, según el artículo 11 de la Ley N° 26790, de Modernización de la Seguridad Social en Salud (LMSSS).
14. Población asegurada acreditada. Población que cumple con los criterios de acreditación establecidos normativamente para acceder a los servicios y prestaciones que otorga EsSalud. En realidad está incluida dentro de la población asegurada activa.

# Modelos de econometría espacial para la lucha contra la delincuencia en el Perú: un enfoque de optimización en tiempo real

Rafael Caparó<sup>1</sup>, Elizabeth Pari<sup>2</sup>

## Resumen

El grado de correlación espacial mostrado entre las regiones evidencia que existe un efecto contagio a nivel de delincuencia sobre todo en las regiones periféricas de Lima, se ha encontrado una relación positiva que avanza de región en región, de distrito en distrito, con tendencias positivas y con claras dependencias con los niveles de ingresos y densidad poblacional, sobre todo en los distritos que se encuentran en los conos, confirmando que el modelo puede contrastar la realidad y predecir escenarios futuros en cuanto a niveles de delincuencia. Si bien la tecnología relacionada al almacenamiento de información es reducida, se presenta una alternativa para la recolección de datos y algunos códigos para almacenar grandes volúmenes de datos, sobre todo los relacionados a las posiciones georreferenciadas de zonas de alta delincuencia, los cuales han servido para encontrar patrones de expansión regional de actos delictivos, monitorear los mismos y proponer alternativas de mejora.

**Palabras clave:** Correlación espacial, delincuencia, posiciones georreferenciadas.

## Introducción

Los índices de criminalidad y la sensación de inseguridad de la población son factores que atentan contra la calidad de vida de las personas, hay una constante preocupación de las autoridades por conocer los factores que conducen a los individuos a cometer

<sup>1</sup> rcaparoc@uni.edu.pe

<sup>2</sup> lizpari.h@gmail.com



un delito<sup>1</sup>. Cerca de 10 millones de peruanos, la tercera parte de la población ha sido víctima de los delincuentes. En el Perú y especialmente en las grandes ciudades la delincuencia es cada vez mayor, hurtos menores, balaceras y asaltos en buses entre otros actos delictivos son ejemplo de momento extremo que estamos viviendo y que por mucho repercute en el bienestar de la población.

Esta investigación muestra factores que influyen el incremento de la delincuencia en Lima y su generalización para todo el territorio peruano. Consideramos un análisis espacial y el uso de nuevas técnicas de econometría espacial para explicar el fenómeno de difusión de la delincuencia en diferentes territorios. Esta investigación hace énfasis también en la correlación espacial de los actos delictivos que pueden llevar a contagios entre regiones. La investigación busca ser pionera en el uso de instrumentos espaciales en tiempo real para poder diseñar medidas proactivas que apoyen la detección de actos delictivos y la mitigación de los mismos.

La importancia de poder cuantificar con datos reales problemas sociales como la delincuencia se vuelve un factor importante entre los que hacemos investigación, de esta manera este documento se centra en el desarrollo de un modelo de ingeniería económica basado en modelos de econometría espacial para medir, monitorear y mitigar problemas de delincuencia en el corto, mediano y largo plazos. La diferencia de este modelo frente a otros radica en las expectativas tecnológicas que pueden aplicarse para una optimización de los recursos, en cuanto a temas regionales de espacio y ubicación contra la delincuencia. Las metodologías matemáticas de econometría espacial ya se utilizan en otros países y muestra resultados favorables.

Esta investigación se concentra en encontrar modelos con información histórica, capaces de capturar patrones de actos delictivos de manera espacial georreferenciada y basada en la econometría espacial que ayuden a la lucha contra la delincuencia. En el corto plazo se optimiza la distribución de recursos en función a focos o centros de delincuencia (ubicados de manera exacta y georreferenciada, gracias al avance tecnológico) para evitar un efecto contagio hacia otras regiones (distritos, barrios, localidades, etc.); este efecto contagio está relacionado con la correlación espacial de actos delictivos que previamente se debe encontrar.

Para el mediano y largo plazo se pretende desarrollar modelos con data más antigua y encontrar las variables más significativas que determinan la delincuencia o que la explican. De esta manera se puede hacer un modelo basado en datos históricos que expliquen los causantes históricos de los actos delictivos de una población en particular, como por ejemplo, falta de acceso a la educación, pobreza, características familiares, ingreso familiar, etc. De este modo, un modelo a largo plazo converge en recomendaciones para políticas de Estado.

<sup>1</sup> Análisis econométrico de los determinantes de la criminalidad

## ¿Qué dice la teoría?

A continuación, se presenta un detalle de los principales documentos que resumen experiencias similares a las que se viven en Perú en cuanto a la lucha contra la delincuencia haciendo análisis espacial, económico, intrínseco y extrínseco del acto delictivo en sí.

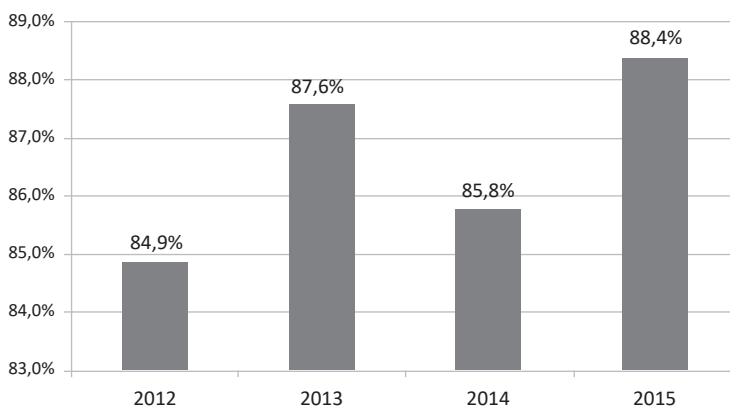
Mediante un análisis espacial toma la variabilidad del hurto y algunas variables que pueden explicarlo. Una de las conclusiones más relevantes, es que la estabilidad residencial tiene un efecto positivo en la tasa de hurto, es decir, que los vecindarios más estables presentan bajas tasas de hurto y robo. Adicionalmente, el aumento en los niveles de capital social obedece a menores tasas de robo (Martin, 2002).

Para el caso de Colombia, la investigación buscó analizar las manifestaciones de violencia y actividades criminales con el conflicto armado (1946-1962) e hizo un modelo de regresión espacial autorregresivo, que toma como variables de dinámica espacial y temporal de los actores armados con las variables sociales, de justicia y narcotráfico (Sánchez, Díaz y Formisano, 2003). Los resultados para el caso colombiano indican que la actividad criminal del secuestro genera efectos importantes de persistencia temporal y de difusión contagiosa, y la existencia de una relación positiva y significativa entre el conflicto y la violencia homicida. Por ello, el concepto de difusión contagiosa fue corroborado al encontrar que la tasa de homicidios del departamento vecino tiene un efecto positivo en la tasa de homicidios local (Barón, Sandoval y Martínez, 2008). En pocas palabras, no muestra de forma contundente las muertes relacionadas directamente con el conflicto.

Tudela (2012) muestra que la correlación espacial determina la interacción y dependencia espacial y temporal de las actividades delictivas en los departamentos de Bolivia. Analiza variables como delitos contra la integridad corporal y la salud, homicidios, hurtos, robos, atracos o robos agravados, violaciones, estupros y abusos deshonestos y busca explicar estos hechos delictivos mediante la realidad económica, representada por las variables Producto Interno Bruto y Producto Interno Bruto Per cápita; también revisa variables sociales como el número de operativos realizados y la cantidad de droga incautada; y finalmente variables ambientales como la temperatura. La investigación toma como referencia estos datos y los analiza a través de modelos econométricos espaciales de panel para mostrar los impactos de las variables descritas en las actividades delictivas y sus niveles incrementales (Tudela, 2012).

Cada país legisla en su constitución y sus códigos penales el tratamiento legal de la delincuencia y los de actos delictivos. En este trabajo se trata de llegar a los niveles de delincuencia más graves a partir del análisis de los ilícitos más simples como los asaltos a mano armada o los robos de celulares. Sin embargo, algunos datos que se debe tener en cuenta para plantear una buena investigación, se han tomado de instituciones militares y del Ministerio del Interior del Perú. El gráfico 1 es un ejemplo de esta afirmación.

**Gráfico 1**  
**Percepción de la inseguridad**



Fuente: Observatorio Nacional de la Seguridad Ciudadana, MININTER.

El gráfico representa claramente lo poco seguros que nos sentimos en el caso urbano del Perú. Además, en un ejercicio de desagregación de los datos podemos ver que son las grandes ciudades peruanas en donde esta “sensación” es aún mayor.

Dentro de los diferentes actos ilícitos penales que son mayormente mencionados por los peruanos podemos encontrar lo siguiente:

**Tabla 1**  
**Percepción de inseguridad por tipo de delito**

Tipo de hecho delictivo	Enero-Junio 2013	Enero-Junio 2014	Enero-Junio 2015
Robo de dinero, celular	78.9	77.6	77.5
Robo de vivienda	70.3	70.1	74.9
Robo de auto	46.4	44.3	39.1
Amenazas	38.1	38	38.8
Extorsión	-	7.9	20.3
Maltrato y ofensa sexual	14.5	13.6	13.1
Secuestro	13.2	12.8	14.1

Fuente: INEI. Estadísticas de Seguridad Ciudadana, aplicado a población urbana mayor de 15 años.

Dentro de las justificaciones de realizar esta investigación se encuentran: la necesidad de las regiones de disminuir los índices de tasas de delincuencia dado que afecta el crecimiento del bienestar social en sus territorios; asimismo, poder detectar las variables globales que podrían afectar las decisiones de una persona a ser más

proclive a que pueda cometer actos delictivos. Es en estas variables donde deberían tener más cuidado los gobernantes, de modo que la toma de sus decisiones no genere una carga de preocupación a la población.

Es relevante resaltar la importancia de la prevención de cualquier acto ilícito que altere las normas que pueda relacionar o involucrar a las empresas en irregularidades, delitos contra el Estado o entidades públicas o privadas, por lo que se debe contar con suficientes medios de información fidedigna que permitan atender las exigencias del entorno en materia de toma de decisiones políticas y la implementación de nuevas herramientas en el marco local.

Este documento tiene como objetivo presentar las variables cualitativas que se pueden cuantificar en cuanto a la toma de decisión para una variación de tasa de delictividad en la región, y mostrar así que la misma se afecte a través del tiempo siguiendo un “contagio colectivo”.

De la misma manera, esta investigación pretende medir experiencias que puedan ser usadas como referencia en cuanto a la detección del incremento de delincuencia; busca variables que lo afecten y toma como referencia la tasa de desempleo, la distribución del territorio y el ingreso promedio. Se construye así un modelo para la detección de variación de tasas de delictividad por parte de la población, de tal forma controlar este riesgo y disminuir en el posible impacto negativo que se pueda reflejar en una toma de decisión de entes negativos o de contagio.

En cuanto a la detección y prevención se buscará combinar con otras técnicas para encontrar los patrones de comportamientos sospechosos, así en un futuro trabajar con tecnologías que propicien esta ayuda a las instituciones a predecir el crimen. Tales tecnologías son herramientas que se prestan para desarrollar la inteligencia artificial.

El ámbito espacial de este trabajo de grado es el territorio peruano, particularmente en el departamento de Lima, además de datos obtenidos a nivel nacional que muestren ser vulnerables a los impactos de decisiones políticas. Asimismo, este trabajo se limita a la selección de variables explicativas para la toma de decisiones a nivel cuantitativo, dado que no se cuenta con los recursos necesarios para una investigación cualitativa de las diversas reacciones de la población a tener tendencia del impacto que puede causar una variación de cambios políticos, geográficos, etc. No se cuenta con datos en tiempo real tales que nos podría proporcionar una gran base de datos (big data).

## **Factores que influyen en la generación de actos delictivos**

La violencia puede ser explicada mediante factores biológicos y sociales. Desde la perspectiva biológica las características propias del individuo lo impulsan a cometer actos delictivos. Por otro lado, los factores sociológicos dan importancia predominante a los factores externos, según esto, los agentes serán buenos o malos conforme al ambiente en el cual vivan y se desarrollen (Flores, 2018).

Con relación a la delincuencia, Middendorff (1964) señala que el ambiente local puede ejercer un fuerte influjo sobre la extensión y clases de la criminalidad. En este marco existen las denominadas “áreas de desorganización social”, las cuales están caracterizadas por una prosperidad deteriorada, existencia de grupos al margen de la Ley y heterogeneidad cultural y social (Tudela, 2011).

La importancia de la desigualdad en los niveles de inseguridad de las ciudades, si bien puede sonar una idea intuitiva, ha sido corroborada desde estudios económicos (Zárate, Aragón y Morel, 2013). Así, los autores Fajnzylber, Lederman y Loayza (2002) encontraban que:

Las tasas de criminalidad y de inequidad están positivamente correlacionadas (dentro de cada país y, de manera especial, entre países), y parece que esta correlación está reflejando una relación causal que va desde la inequidad a las tasas de criminalidad aún después de haber sido controlada por otras posibles determinantes (2002a: 26).

Por otro lado, Formisano (2002) agrupa tres tendencias, de las cuales la primera explica la violencia como resultado de factores socioeconómicos (pobreza o desigualdad). Otra de las tendencias es denominada “Economía del Crimen”; la cual ha vivido un desarrollo reciente a nivel mundial y nacional, y establece como principal resultado que el gasto asociado a justicia y seguridad disuade a los individuos de cometer delitos, ya que aumentan los costos coligados a delinquir. Por último está la teoría del “desorden”, la cual afirma que el desorden existente en un barrio genera mayor número de delitos. Por ejemplo, para el caso de Nueva York, Corman (2000) afirma que la presencia de policía y las capturas de criminales fueron factores disuasivos del crimen y redujeron la tasa de homicidios a finales de los años 90. La evidencia empírica en Estados Unidos también confirma la relación causal y positiva entre la posesión de armas de fuego y la tasa de homicidios (Corman y Mocan, 2000).

Entre otras teorías, la teoría del desorden urbano o “Broken Window” sostiene que existe una relación de causalidad positiva y directa entre las características de las zonas donde se cometen los crímenes y los niveles de criminalidad; de esta manera, las zonas desordenadas (con presencia de prostitución, expendio y consumo de drogas, pandillas y vandalismo, entre otras) favorecen la implantación y el desarrollo de todo tipo de actos violentos (Formisano, 2002). En ese sentido, Kelling y Willson (1996), realizan una analogía al afirmar que igual que en la violencia, en un edificio, si se rompe una de sus ventanas y esta se deja sin reparar, las demás ventanas serán prontamente rotas.

El trabajo de Llorente, Escobedo, Echandía y Rubio (2000) asocia las variables de muertes violentas con la presencia de prostitución, consumo de drogas y zonas de venta de drogas, tal relación fue establecida mediante la observación y superposición de mapas de la tasa de homicidios y las distintas variables.

## Econometría espacial: El estadístico I de Morán

Una de las formas de medir la autocorrelación espacial es el estadístico I de Morán. Este estadístico permite probar dos hipótesis:

1. Si la variable dependiente está concentrada espacialmente,
2. Medir los cambios temporales en los niveles de concentración.

La lógica del procedimiento es que un aumento temporal en la magnitud de los coeficientes indicaría un proceso de divergencia regional, mientras que una disminución indicaría evidencia de convergencia regional (Marín, 2003).

El estimador del coeficiente de I de Moran es:

$$I = \frac{[Z']^T [W \otimes I] [Z']}{[Z']^T [Z']}$$

Donde el valor I puede tomar diferentes estructuras. Estas estructuras se determinan de acuerdo a la estructura de vecindades. De estas depende la especificación de la matriz  $W^*$  la cual es determinante en la construcción del modelo de regresión espacio temporal.

## Perú productor de droga

Según refiere un artículo del diario Gestión, publicado el 13 de julio de 2016, Naciones Unidas señala al Perú como el segundo productor de cocaína en el mundo:

*Durante el 2015 se sembraron 40,300 hectáreas de cicales en Perú, lo que significó una reducción de 6.1%, en comparación con 2014. La cifra mantiene una tendencia a la baja que empezó en 2011 cuando se detectó que se habían sembrado 62,500 hectáreas de hojas de coca. La mayor cantidad de cicales están en dos zonas: el valle de los ríos Apurímac, Ene y Mantaro y el valle de La Convención-Lares, ambas en el sureste. El informe indica que 87,304 toneladas de hojas de coca van al narcotráfico, mientras que apenas 9,000 toneladas son usadas para el masticado de la hoja de coca u otras formas de consumo tradicional. El precio del kilo de cocaína en Perú disminuyó su valor en 2015 retrocediendo hasta 1,133 dólares, luego que en 2014 alcanzó un promedio de 1,178 dólares. En Perú el precio del clorhidrato alcanzó su valor máximo en 2013 cuando trepó hasta 1,310 dólares el kilo, según la ONU.*

Se calculó la correlación entre la cantidad de droga decomisada por distritos a nivel nacional si tiene una implicancia en cuanto al nivel de estudios por región. A lo largo del tiempo los departamentos de Ayacucho, Cusco, Huánuco, La Libertad, Piura y Lima arrojan cifras alarmantes de distribución de droga. En lugares donde han habido incautaciones se observa que no implica mucho el nivel educativo para que estas personas encontradas decidan dedicarse a esta actividad ilícita.

Gráfico 3  
Delitos genéricos por departamento

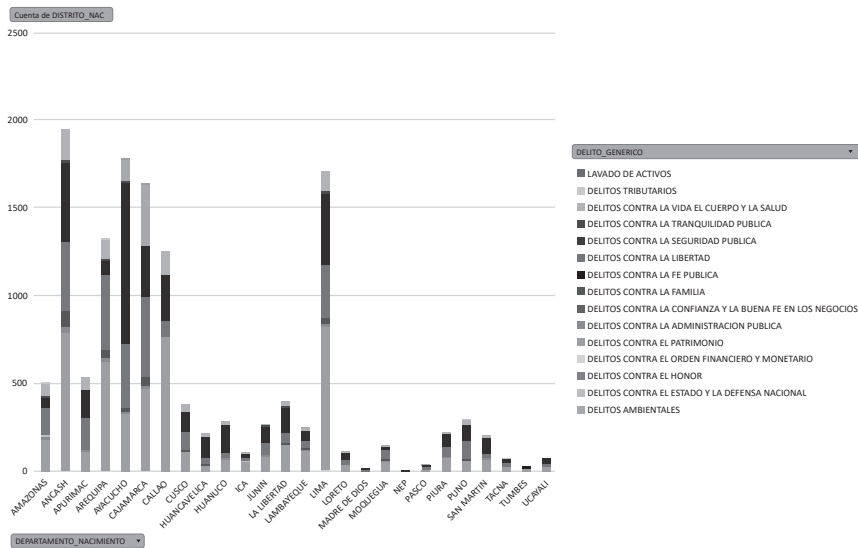


Gráfico 4  
Delitos genéricos por departamento

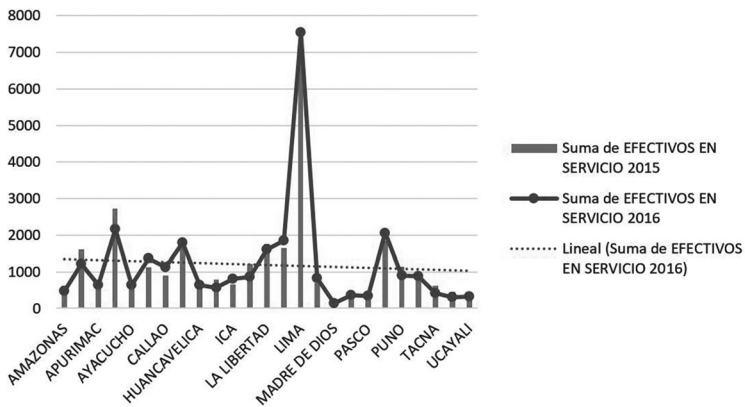


Tabla 1  
Droga decomisada a nivel nacional

Departamento	Droga decomisada (kg) 2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	TOTAL
Amazonas	16	13	6	66	40	5	39	139	52	125	42	48	590
Áncash	47	132	121	181	656	211	1547	945	552	220	328	243	5 184
Apurímac	325	177	138	287	454	137	822	399	575	226	462	99	4 102
Arequipa	14	-	38	210	72	21	116	1352	437	426	503	1	4 187
Ayacucho	1 588	841	974	1330	2144	1774	3574	3916	2378	2215	4275	3796	28 807
Cajamarca	479	98	55	40	11	7	79	28	40	54	64	168	1 123
Cusco	454	405	264	332	197	639	682	1718	2532	914	2024	2728	12 888
Huancavelica	265	104	152	-	16	106	31	301	60	-	21	273	1 329
Huánuco	806	446	976	793	1208	2752	1756	1815	3178	2 971	3303	34462	54 468
Ica	260	146	63	207	1053	1529	1654	1321	1349	808	615	571	9 776
Junín	1 045	2543	1257	953	2899	1967	612	1922	4157	4666	3316	10034	35 371
La Libertad	142	119	137	160	426	290	1006	789	499	279	7551	53071	64 470
Lambayeque	105	346	104	109	1191	135	14)	71	233	280	101	493	3 307
Lima 1/	2917	8677	9928	7220	16812	9501	16272	7674	12309	5553	7343	10576	114 792
Loreto	385	562	1076	611	725	262	610	599	916	710	458	551	7 465
Madre de Dios	100	104	259	471	306	181	238	282	286	75	105	41	2 449
Moquegua	14	6	3	354	46	89	2450	829	231	226	55	19	4 323
Pasco	192	17	41	104	122	97	136	-	1377	852	421	1654	5 013
Piura	5 604	1317	478	1061	520	742	252	1197	1935	4761	353	1689	19 909
Puno	134	281	152	619	763	976	836	1574	585	450	370	284	7 024
San Martín	247	332	139	165	524	539	301	220	188	748	1010	487	4 900
Tacna	152	1128	88	178	61	102	128	197	179	228	235	312	2 988
Tumbes	88	154	104	43	139	81	271	94	788	35	73	54	1 924
Ucayali	235	71	166	190	238	699	288	403	578	1173	814	1809	6 664

Pero si la actividad ilícita de comercialización y producción de droga influye más en estos tipos de delitos que, efectivamente, viene del mercado negro, tales como aumento de armas de fuego en homicidios calificados, asesinato, debido a esta actividad se ve una mayor proporción ocurrida en la vivienda de la víctima, mayormente es efectuada por un desconocido, en similar proporción a causas como venganza y/o delincuencia organizada. Otros lugares donde la droga ha impactado son los delitos asociados a lugares como vehículos en la vía pública.

El la tabla 2 se muestra las implicancias de la correlación de la droga con los delitos asociados en conjunto de los lugares más perpetrados así como las causas y/o motivo del delito y la proveniencia de la relación de la víctima con el delincuente, mayormente son realizados por desconocidos, seguidos de que la gran mayoría de las víctimas conoce al victimario, otro punto resaltante es que



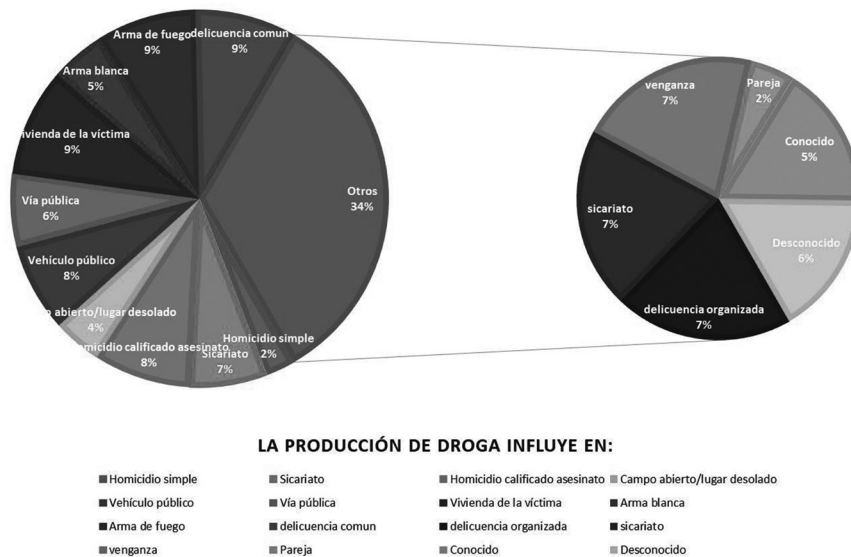
se realizan como lugares más comunes siendo dentro de la vivienda y el aumento del uso de arma de fuego.

Tabla 2  
Correlación de la droga (Año 2015)

Homicidio simple	0.18	Arma de fuego	0.60
Sicariato	0.47	Delincuencia común	0.61
Homicidio calificado asesinato	0.59	Delincuencia organizada	0.48
Campo abierto/lugar desolado	0.30	Sicariato	0.49
Vehículo público	0.53	Venganza	0.49
Vía pública	0.44	Pareja	0.12
Vivienda de la víctima	0.65	Conocido	0.39
Arma blanca	0.33	Desconocido	0.39

Fuente INEI, elaboración propia

Gráfico 5  
Influencia de la producción de droga

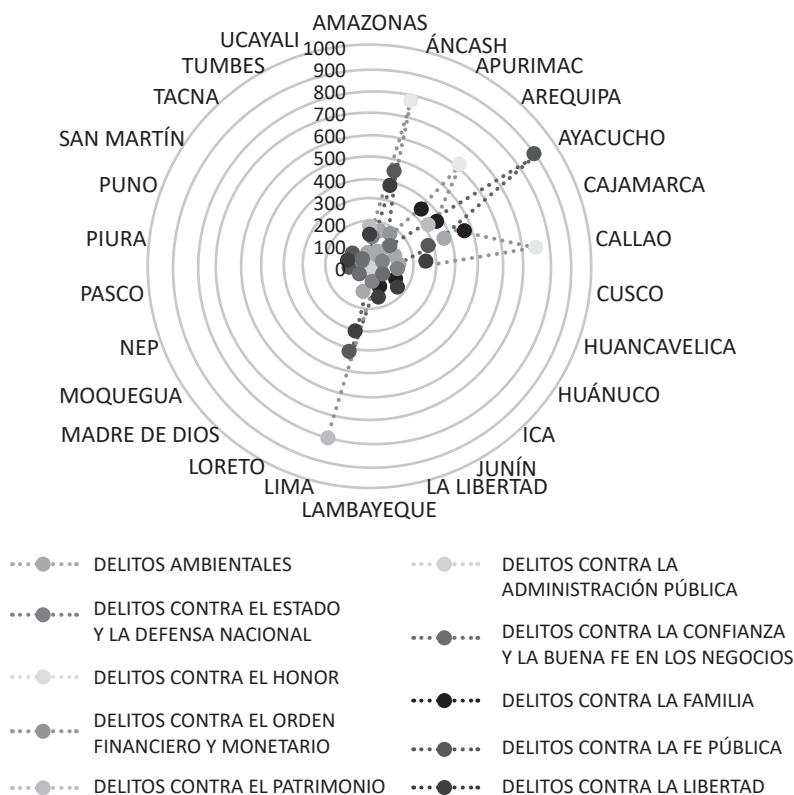


## Delitos por distritos

De una cantidad de 13 860 datos recopilados de las distintas intendencias policiales, datos recopilados del INEI de 985 distritos a nivel del Perú, en el siguiente cuadro se puede observar que los departamentos con mayor cifras de denuncias son Ayacucho, Lima y Callao, seguido de Cusco; en mayor medida se cometen delitos contra el patrimonio.

Al nivel departamental, los delitos más cometidos en la ciudad de Lima son el homicidio simple y el calificado, así como la emoción violenta. Se busca saber acerca del factor de mayor delincuencia en el país, por ello se tiene como factor el tráfico ilícito de drogas, actividad que genera grandes sumas de dinero.

Gráfico 1  
Delitos por departamentos

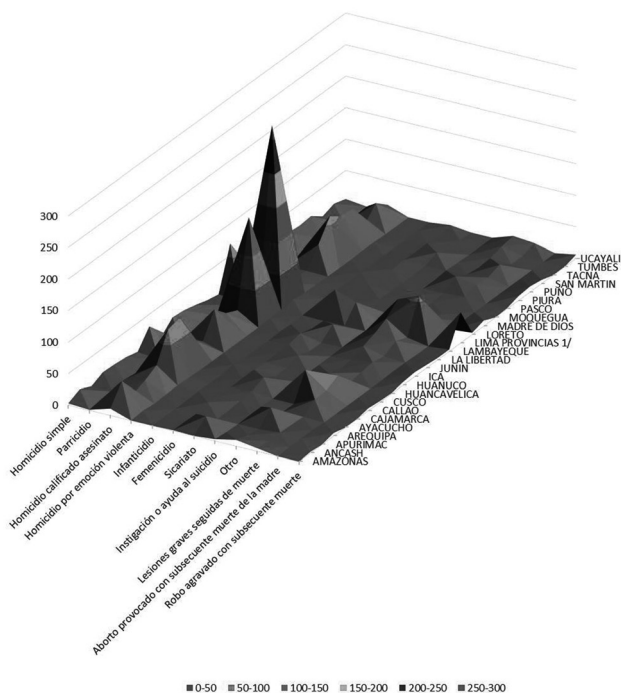


Fuente INEI, elaboración propia.

En el gráfico siguiente se visualiza que hay un mayor porcentaje en delitos contra la vida y salud, así como delitos contra la seguridad pública, que tienen la mayor cantidad de delitos en la ciudad Ancash para el año 2015.

Según una encuesta del INEI, en las comisarías se aprecia la disminución de efectivos por cantidad de habitantes; y aunque la tendencias es a la baja, en el Callao se ve el aumento de efectivos en las comisarías para el año 2016.

Gráfico 2  
Delitos por departamentos



Fuente INEI, elaboración propia

## El método

Este trabajo es de tipo descriptivo-analítico, ya que permite ordenar el resultado de las observaciones, con el fin de poder detectar las variables que resulten más representativas. La metodología comprende los siguientes pasos:

1. Análisis estadístico de datos y comparación por regiones.
2. Identificación de correlaciones espaciales.
3. Construcción, análisis, diagnóstico e interpretación de un modelo econométrico espacial.

Para ello se disponen de las siguientes variables de actividades delictivas:

- Delitos contra la integridad corporal y la salud de las personas.
- Homicidios
- Hurto
- Robo
- Robo agravado (atracó)
- Violación, estupro y abuso deshonesto.

Las mismas que se desean explicar con las siguientes variables:

- Variables económicas
  - Ingreso
  - Tasa de desempleo
- Variables sociales
  - Coeficiente de Gini

Con esto se pretende responder a preguntas tales como: ¿qué indicador podrá reflejar de manera tácita si alguna zona afectada sufre problemas, como por ejemplo, de inseguridad ciudadana? Según lo explican algunos hechos, hay más delitos en la zona donde uno vive que otras zonas. Se quiere llevar a un plano más cuantitativo el análisis, partiendo del plano cualitativo, aunque su valorización puede ser cuestionable y debatible por la comunidad, redes sociales y medios de comunicación.

Si bien existe monitoreo en zonas urbanas donde se puede notar que la percepción de inseguridad no ha desaparecido, sino que incluso se ha incrementado en cifras altas, esa razón acarrea una gran variedad de problemas concernientes al incremento de inseguridad, como asaltos a mano armada, hurtos, secuestros, extorsiones, cobro de cupos, violencia familiar, robo de autopartes, etc.

Hay que destacar que la composición de esta percepción varía mucho en función de variables como edad, sexo, nivel socioeconómico o ubicación geográfica. Ahora, ¿cuánto de esa percepción se llega a “materializar” en la comisión de un ilícito contabilizable para fines estadísticos? En siguientes informes definiremos lo que la literatura pertinente conoce como victimización y cómo daremos el salto al plano cuantitativo al dar a conocer este indicador como reflejo de la Inseguridad ciudadana de una región o comunidad.

Las herramientas que posteriormente presentaremos y usaremos tienen como materia prima datos, representados por variables, con lo que surge la pregunta ¿Qué variables pertinentes al tema de la inseguridad ciudadana podríamos identificar como útiles para nuestros propósitos?

**Número de delitos por distrito:** viene a ser aquel que captura el efecto que deseamos explicar, la inseguridad ciudadana. Se define como el número de ilícitos denunciados en la comisaría del sector. Intuimos que a mayor número de delitos en una zona determinada, mayor inseguridad ciudadana sufre. Para ello recopilamos la data del número de delitos cometidos en los 43 distritos de Lima (los 6 del Callao inclusive) para los años 2014 y 2015.

**Desempleo:** Se hizo el análisis con los ratios de desempleo para los años 2014 y 2015, donde se observa que existe mayor desempleo a comparación del año anterior. Los distritos con mayor ratio de desempleo son San Bartolo y Magdalena Vieja para el año 2014 y 2015, respectivamente; mientras que el distrito con menor ratio de desempleo es Pucusana (0.01) que se mantiene en segundo lugar para el siguiente año; el distrito con menor ratio de desempleo para el 2015 es Independencia (0.15). Los distritos que se mantienen en el promedio para el año 2014 son Punta Hermosa

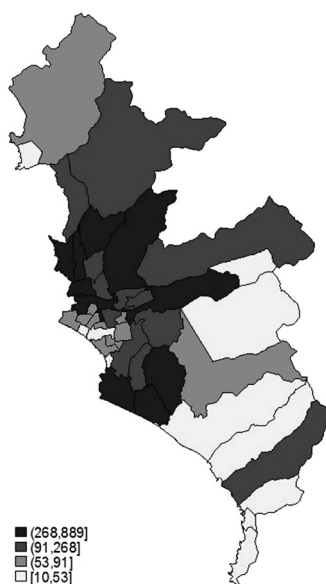
(1.27), Chaclacayo (1.32) y Villa María del Triunfo (1.33) y para el año 2015 se mantienen los distritos de Carabaylo (1.36), Villa María del Triunfo (1.36) y San Miguel (1.37). Se tomó como datos personas en edad a trabajar y personas desempleadas. También puede explicarse el aumento de la variable de desempleo por el incremento de la (PEA) y la población por distrito.

**Ingreso promedio:** De los datos obtenidos, se hizo el análisis del ingreso promedio por distritos con datos de personas que se encuentran laborando. Se debe tener en cuenta que algunos hogares reciben mensualmente el mismo ingreso mientras otros hogares realizan trabajos independientes.

Se observa que existe una mayor cantidad de personas que tienen de ingreso una aproximación de S/. 1116 para el año 2014. Los distritos con esta aproximación al ingreso promedio son San Martín de Porres, Comas y La Victoria. Asimismo, para el año 2015 el ingreso promedio subió a S/. 1333, aproximadamente, y la mediana a nivel distrital fue de S/. 1239, entre estos distritos se encuentran Chorrillos, El Agustino y Los Olivos. Los distritos con menor y mayor ingreso para el año 2014 son Pucusana y San Isidro, respectivamente; el año 2015 Cieneguilla y San Borja fueron los distritos con menor y mayor ingreso. En comparación del año 2014, el año 2015 se observó un aumento del ingreso promedio en un 2,56% de crecimiento a nivel distrital.

Se ha observado que algunos distritos como Cieneguilla bajaron considerablemente su ingreso promedio y esto puede explicarse por factores exógenos.

Figura 1  
Cantidad de delitos en Lima Metropolitana (Año 2015)



Fuente: INEI.

## Metodología para el trabajo de los datos: big data y datos en tiempo real

A continuación, hacemos un resumen de los principales pasos para procesar la gran cantidad de datos, su almacenaje y su difusión:

### Convirtiendo archivos shapes a stata

Para iniciar el análisis espacial es necesario tener información georreferenciada. Un problema habitual es la obtención de dicha información. Sin embargo, esta restricción se ha vuelto menos importante a medida que diferentes sitios webs han ido ofreciendo archivos georreferenciados de manera gratuita. Entre los sitios gratuitos puede mencionarse Shapfiles Perú, que provee archivos de mapas base y coordenadas para los distritos de Lima.

El formato shape se ha convertido en el tipo de archivo más utilizado para el intercambio de información geográfica. Un archivo shape es un formato vectorial que almacena la localización de elementos geográficos y sus atributos. La diferencia con otro tipo de archivo es que es un formato multiarchivo: conjunto relacionado de ficheros informáticos. El número mínimo y habitual es de tres ficheros con las siguientes extensiones:

- **.shp** es el archivo que almacena las entidades geométricas de los objetos.
- **.shx** es el archivo que almacena el índice de las entidades geométricas.
- **.dbf** es la base de datos, en formato dBASE, donde se almacena la información de los atributos de los objetos.

El primer paso es cargar la información georreferenciada mediante el siguiente comando:

- `shp2dta using BAS_LIM_DISTritos, database(datos_shp) coordinates(coord) /// genid(id) genc(c)`

Este comando lee el archivo shape "BAS\_LIM\_DISTritos", generando dos archivos: "datos\_shp.dta" y "coord.dta". En el archivo de datos, "datos\_shp", se genera adicionalmente una variable identificadora de cada unidad geográfica, **id**, y los centroides, puntos que definen el centro geométrico de los polígonos, identificados por la latitud (**y\_c**) y la longitud (**x\_c**).

- `use datos_shp, clear`
- `describe`

El archivo `datos_shp.dta` contiene la información sobre los atributos del archivo `dbf`. El archivo de atributos tiene únicamente las variables que identifican a cada `BAS_LIM_DISTritos`, el nombre del objeto, y las coordenadas geográficas.

Pisati (2008) provee el comando **spmap** que permite obtener un gráfico del mapa con los límites de cada área, es decir, el mapa base.

- spmap using coord, id(\_ID) note("Lima, 2015")

Figura 2  
**División política de la provincia de Lima**



Lima, 2015

## Especificación del modelo

Dentro de la amplia variedad de modelos espaciales deben destacarse dos estructuras que con más frecuencia se utilizan en econometría espacial: los llamados modelos de retardo espacial (*Modelos de comunicación o contagio*) y los modelos del error espacial o de *comparación*. Los primeros, los modelos de retardo espacial, recogen la estructura de dependencia espacial del proceso mediante la inclusión de un retardo espacial como factor explicativo de la variable endógena, con la siguiente especificación:

$$Y = \rho WY + X\beta + e$$

$$e \sim N(0, \sigma^2 I)$$

Donde como es habitual,  $Y$  es un vector columna  $n \times 1$ ,  $X$  es una matriz  $k \times n$  que recoge una serie de variables exógenas (donde opcionalmente se puede incluir una columna de unos para el término constante),  $W$  es una matriz de conexiones  $n \times n$  exógena que define la estructura de vecindades,  $WY$  es el retardo espacial de la variable  $Y$ .

En este modelo la relación espacial se da por medio del denominado contagio donde los choques exógenos se propagan entre las unidades vecinas. La segunda especificación introduce la estructura de dependencia espacial en el término de perturbación del modelo, como se muestra a continuación:

$$Y=X\beta+u$$

$$u=\rho Wu+e$$

$$e\sim N(0,\sigma^2 I)$$

Como en el caso anterior, Y es un vector columna nx1, X es una matriz kxn que recoge una serie de variables exógenas, W es una matriz de conexiones nxn exógena que define la estructura de vecindades y es el parámetro de dependencia espacial.

En este análisis nos centraremos en la estimación clásica donde nuestras variables recogidas son el ingreso promedio y la tasa de desempleo observadas en cada uno de los años desde el 2012 hasta el 2015 acompañadas del término error.

$$\text{Delictividad}=\text{ingreso promedio}\beta_1 + \text{tasa de desempleo}\beta_2 + u$$

## Efectos marginales

Partimos de la estimación del modelo elegido:

- use datos\_shp, clear
- spreg ml delitos2013 delitos2014, dl(W5\_st) id(id)

Efectos Directos e Indirectos según LeSage y Pace (2009)

Efecto Total usando la secuencia propuesta por Drukker et al. (2011)

Predicción (forma reducida - rform) - calculando y estimando

- predict y0

Asumiendo un cambio unitario en la migración neta

- quietly replace delitos2013 = delitos2013 delitos2014+ 1
- predict y1

\* Variación entre predicciones: Impacto total

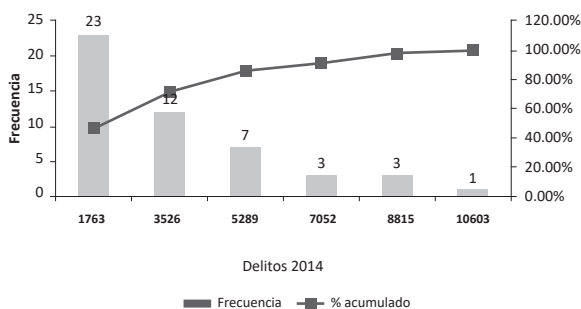
- gen delta\_y = y1-y0
- sum delta\_y



## Datos

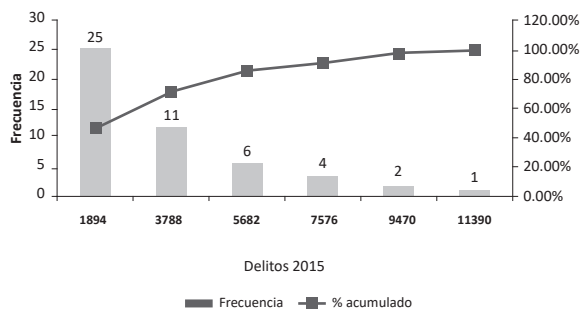
**Número de delitos por distrito:** Viene a ser aquel que captura el efecto que deseamos explicar, la inseguridad ciudadana. Se define como el número de ilícitos denunciados en la comisaría del sector. Intuimos que a mayor número de delitos en una zona determinada, mayor inseguridad ciudadana sufre. Para ello recopilamos la data del número de delitos cometidos en los 43 distritos de Lima (y los 6 del Callao, inclusive) para los años 2014 y 2015, que registra un promedio de 2582 delitos, aproximadamente, a nivel distrital.

Delitos 2014	
Media	2582.367347
Error típico	352.7849185
Mediana	2292
Desviación estándar	2469.494429
Varianza de la muestra	6098402.737
Curtosis	1.54335981
Coefficiente de asimetría	1.327709238
Rango	10578
Mínimo	25
Máximo	10603
Cuenta	49



Fuente: Dirección de estadística de la PNP – MININTER 2014

Delitos 2015	
Media	2621.653061
Error típico	374.1935212
Mediana	1627
Desviación estándar	2619.354648
Varianza de la muestra	6861018.773
Curtosis	1.70179922
Coefficiente de asimetría	1.394332182
Rango	11364
Mínimo	26
Máximo	11390
Cuenta	49

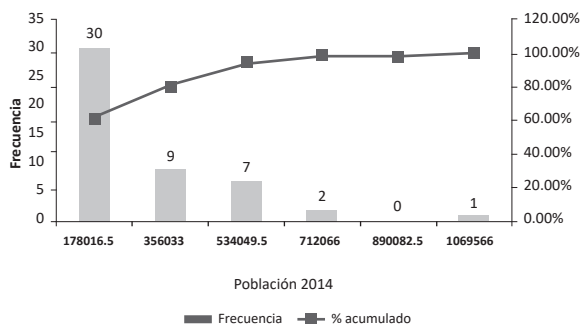


Fuente: Dirección de Estadística de la PNP, MININTER, 2015.

No solo notamos una tendencia al aumento sino que también son recurrentes los distritos en donde se aglomeran denuncias. Además, un distrito sobresale como el de mayor cantidad de delitos: San Juan de Lurigancho: ¿El hecho que sea el distrito más poblado de Lima Metropolitana incide en ello?

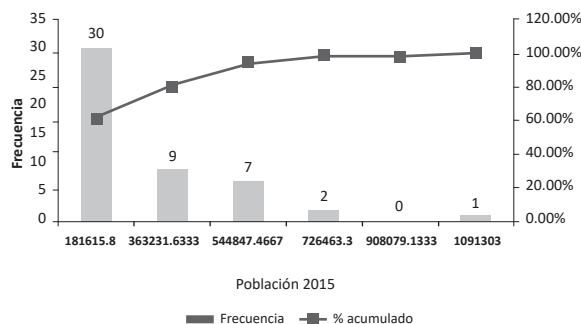
**Población por distrito:** Representa el número de habitantes que tiene un distrito en un lapso de tiempo. Para ello tomamos el número de habitantes en los distritos de Lima para los años 2014 y 2015.

Población 2014	
Media	199014.6327
Error típico	30752.88257
Mediana	111808
Desviación estándar	215270.178
Varianza de la muestra	46341249540
Curtosis	4.520079751
Coefficiente de asimetría	1.858568165
Rango	1068099
Mínimo	1467
Máximo	1069566
Suma	9751717
Cuenta	49



Fuente: INEI. Proyección de población. Junio 2014.

Población 2015	
Media	202137.2357
Error típico	31393.31024
Mediana	111928
Desviación estándar	219753.1717
Varianza de la muestra	48291456467
Curtosis	4.544382669
Coefficiente de asimetría	1.866506719
Rango	1089695
Mínimo	1608
Máximo	1091303
Suma	9904727
Cuenta	49

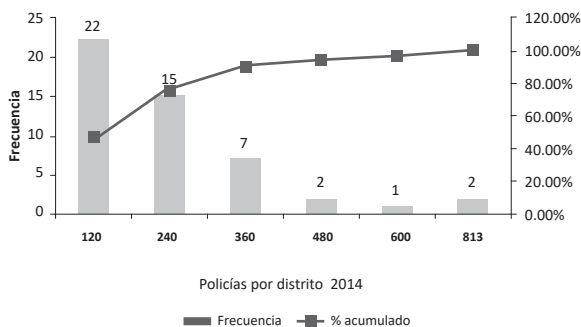


Fuente: INEI. Estimaciones y Proyecciones de población por sexo 2000 -2015

Esta variable población se puede relacionar con densidad poblacional, lo que la involucra con un atributo espacial: ¿a mayor turgurización de un espacio se eleva el número de delitos?

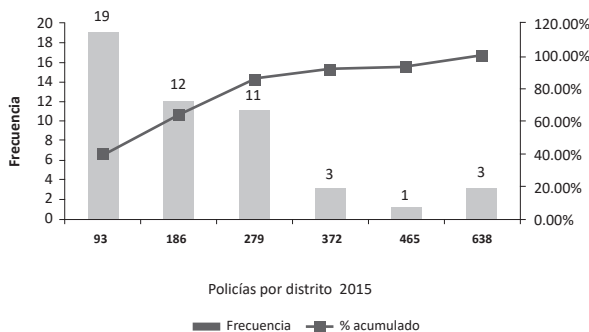
**Número de policías por distrito:** Representa el número de efectivos policiales hábiles en las diferentes comisarías de un distrito. Recopilamos esta data para los años 2014 y 2015 para Lima Metropolitana.

Policías por distrito 2014	
Media	188.795918
Error típico	23.2826462
Mediana	163
Moda	264
Desviación estándar	162.978524
Varianza de la muestra	26561.9991
Curtosis	5.49409048
Coefficiente de asimetría	2.10247323
Rango	791
Mínimo	22
Máximo	813
Suma	9251
Cuenta	49



Fuente: Dirección de Recursos Humanos PNP, 2014

Policías por distrito 2015	
Media	172.367347
Error típico	21.2984281
Mediana	137
Moda	251
Desviación estándar	149.088997
Varianza de la muestra	22227.5289
Curtosis	2.83361467
Coefficiente de asimetría	1.64382527
Rango	614
Mínimo	24
Máximc	638
Suma	8446
Cuenta	49

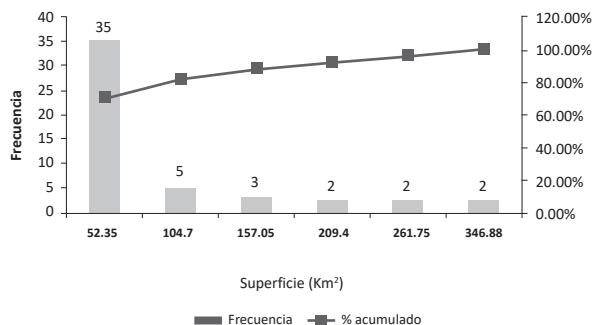


Fuente: Dirección de Recursos Humanos PNP, 2015

Vemos un descenso en el número de efectivos policiales hábiles en Lima, y en particular en varias jurisdicciones, con lo que podemos preguntarnos: ¿un reducido número de efectivos policiales por distrito incide en los índices de inseguridad ciudadana?

**Superficie distrital (Km<sup>2</sup>):** Representa la extensión territorial de cada distrito. Es idónea para practicar análisis espacial de datos. Recopilamos las extensiones de los distritos de Lima Metropolitana.

Superficie (Km <sup>2</sup> )	
Media	57.0208163
Error típico	11.554458
Mediana	21.98
Desviación estándar	80.8812059
Varianza de la muestra	6541.76947
Curtosis	4.08837276
Coefficiente de asimetría	2.10787148
Rango	346.13
Mínimo	0.75
Máximo	346.88
Suma	2794.02
Cuenta	49

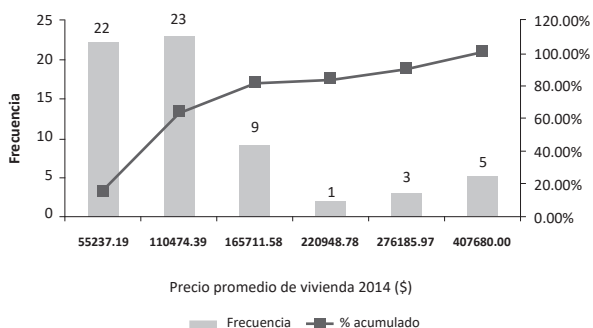


Fuente: INEI. Ubigeo 2010.

Notamos que existe una marcada atomización del espacio en Lima, con una gran cantidad de distritos de pequeña extensión ¿Eso dificulta la unificación de políticas contra el delito? ¿A mayor extensión, más difícil controlar los delitos?

**Precio de vivienda por distrito (\$):** Representa el precio promedio de un departamento en los distritos de Lima Metropolitana. Se ha tomado como referencia un estudio de la Cámara Peruana de la Construcción de 2014.

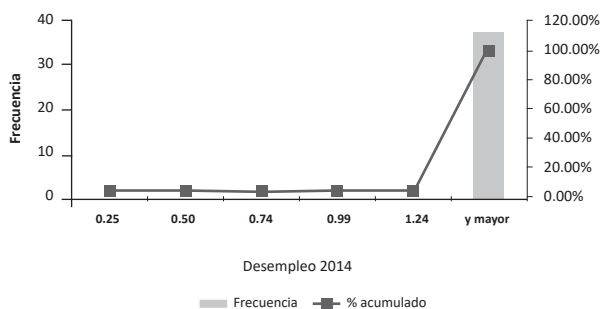
Precio promedio de vivienda 2014 (\$)	
Media	121003.837
Error típico	12165.9436
Mediana	86484
Moda	72560
Desviación estándar	85161.6053
Varianza de la muestra	7252499016
Curtosis	2.2131931
Coefficiente de asimetría	1.63651507
Rango	365198
Mínimo	42482
Máximo	407680
Suma	5929188
Cuenta	49



Fuente: CAPECO 2014.

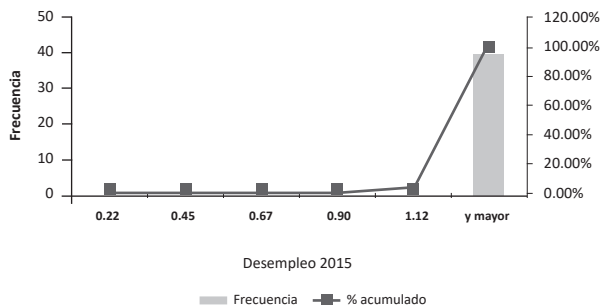
**Desempleo:** Se analizó con los ratios de desempleo para los años 2014 y 2015, y se observó que aumentó el desempleo a comparación del año anterior. Los distritos con mayor ratio de desempleo son San Bartolo (2014) y Magdalena Vieja (2015). El distrito con menor ratio de desempleo en 2014 es Pucusana (0,01) que es segundo en 2015, y el distrito con menor ratio de desempleo en 2015 es Independencia (0,15). Los distritos que se mantienen en el promedio son Punta Hermosa (1,27), Chaclacayo (1,32) y Villa María del Triunfo (1,33) en 2014 y en 2015 se mantienen los distritos de Carabayllo (1,36), Villa María del Triunfo (1,36) y San Miguel (1,37). Se tomó datos de personas en edad de trabajar y de desempleadas. El aumento de la variable de desempleo también se puede explicar por el incremento de la (PEA) y la población por distrito.

Desempleo 2014	
Media	1.31
Error típico	0.06
Mediana	1.41
Moda	1.50
Desviación estándar	0.36
Varianza de la muestra	0.13
Curtosis	9.46
Coefficiente de asimetría	-3.25
Rango	1.49
Mínimo	0.01
Máximo	1.50
Suma	53.64
Cuenta	41.00
Nivel de confianza (95.0%)	0.11



Fuente: INEI, Encuesta Nacional de Hogares (ENAHOG), 2014.

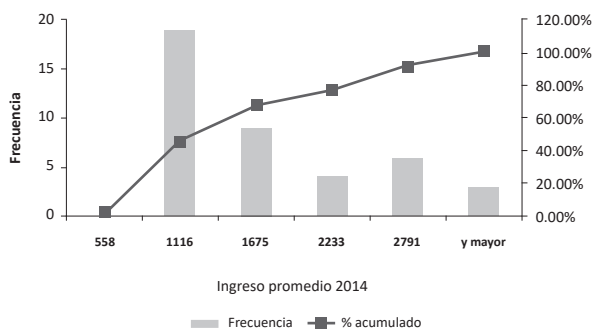
Desempleo 2015	
Media	1.36
Error típico	0.03
Mediana	1.39
Moda	1.41
Desviación estándar	0.21
Varianza de la muestra	0.04
Curtosis	30.14
Coefficiente de asimetría	-5.21
Rango	1.35
Mínimo	0.15
Máximo	1.50
Suma	55.62
Cuenta	41.00
Nivel de confianza (95.0%)	0.07



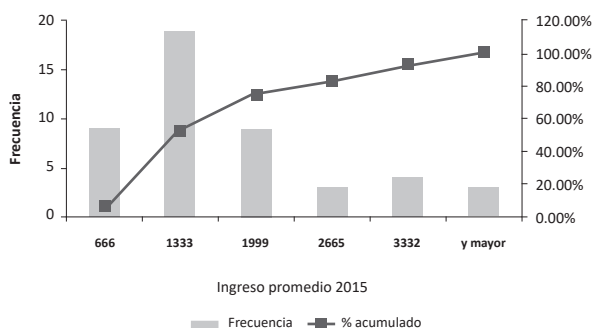
Fuente: INEI, Encuesta Nacional de Hogares (ENAHOG), 2015.

**Ingreso promedio:** Se analizó el ingreso promedio por distritos considerando personas que se hallan laborando. Se debe considerar que algunos hogares reciben mensualmente el mismo ingreso mientras que otros realizan trabajos independientes. Existe una mayor cantidad de personas cuyo ingreso se aproxima a S/. 1116 para el año 2014, los distritos con este ingreso promedio son San Martín de Porres, Comas y La Victoria. Para el año 2015 el ingreso promedio sube a S/. 1333, aproximadamente (mediana a nivel distrital de S/. 1239), entre los distritos con este ingreso tenemos Chorrillos, El Agustino y Los Olivos. Los distritos con menor y mayor ingreso para el año 2014 son, respectivamente, Pucusana y San Isidro; para el año 2015 son Cieneguilla y San Borja. Se observa un aumento del ingreso promedio en un 2.56% a nivel distrital en 2015. No obstante, algunos distritos como Cieneguilla bajaron considerablemente su ingreso promedio, esto puede explicarse por factores exógenos.

ing_prom (2014)	
Media	1532.4
Error típico	130.5
Mediana	1132.0
Moda	-
Desviación estándar	835.4
Varianza de la muestra	697875.1
Curtosis	1.1
Coefficiente de asimetría	1.3
Rango	3349.0
Mínimo	686.0
Máximo	4035.0
Suma	62827.0
Cuenta	41.0
Nivel de confianza (95.0%)	263.68



ing_prom (2015)	
Media	1572.6
Error típico	143.7
Mediana	1239.0
Moda	-
Desviación estándar	919.9
Varianza de la muestra	846286.6
Curtosis	1.4
Coefficiente de asimetría	1.3
Rango	3998.0
Mínimo	398.0
Máximo	4396.0
Suma	64475.0
Cuenta	41.0
Nivel de confianza (95.0%)	290.37



## Resultados

**Manipulación y visualización de datos espaciales:** Ahora es posible visualizar las variables mediante diferentes mapas de coropleto. Este tipo de mapa es un mapa temático en el que las áreas se colorean con una intensidad proporcional al valor que toma una variable cuantitativa (Slocum et al., 2009). Para ello utilizaremos el comando antes presentado, `spmap`.

En `spmap`, un mapa de coropleto es un mapa base cuyos polígonos estarán coloreados de acuerdo al valor tomado de un atributo (variable) que debe ser numérico. La versión actual de `spmap` ofrece seis tipos de coropleto, usaremos principalmente:

- **Cuantiles:** las clases corresponden a los cuantiles de la distribución de la variable, cada clase tendrá aproximadamente el mismo número de polígonos.

- `use datos_shp, clear`

- `format delitos2012 %12.1f`

- `spmap delitos2012 using coord, id(_ID) clmethod(q) title("Cantidad de Delitos ") legend(size(medium) position(7)) fcolor(Blues2) note("Lima Metropolitana, 2012" "Fuente: Inei")`

- `format delitos2013 %12.1f`

- `spmap delitos2013 using coord, id(_ID) clmethod(q) title("Cantidad de Delitos ") legend(size(medium) position(7)) fcolor(Blues2) note("Lima Metropolitana, 2013" "Fuente: Inei")`

```
format delitos2014 %12.1f
```

```
spmap delitos2014 using coord, id(_ID) clmethod(q) title("Cantidad de Delitos ") legend(size(medium) position(7)) fcolor(Blues2) note("Lima Metropolitana, 2014" "Fuente: Inei")
```

```
format delitos2015 %12.1f
```

```
spmap delitos2015 using coord, id(_ID) clmethod(q) title("Cantidad de Delitos ") legend(size(medium) position(7)) fcolor(Blues2) note("Lima Metropolitana, 2015" "Fuente: Inei")
```

Puede combinarse la información de ambas variables por ejemplo:

```
spmap delitos2012 using coord, id(_ID) fcolor(Blues2) point(data(migr_unemp_shp) xcoord(x_c) ///
```

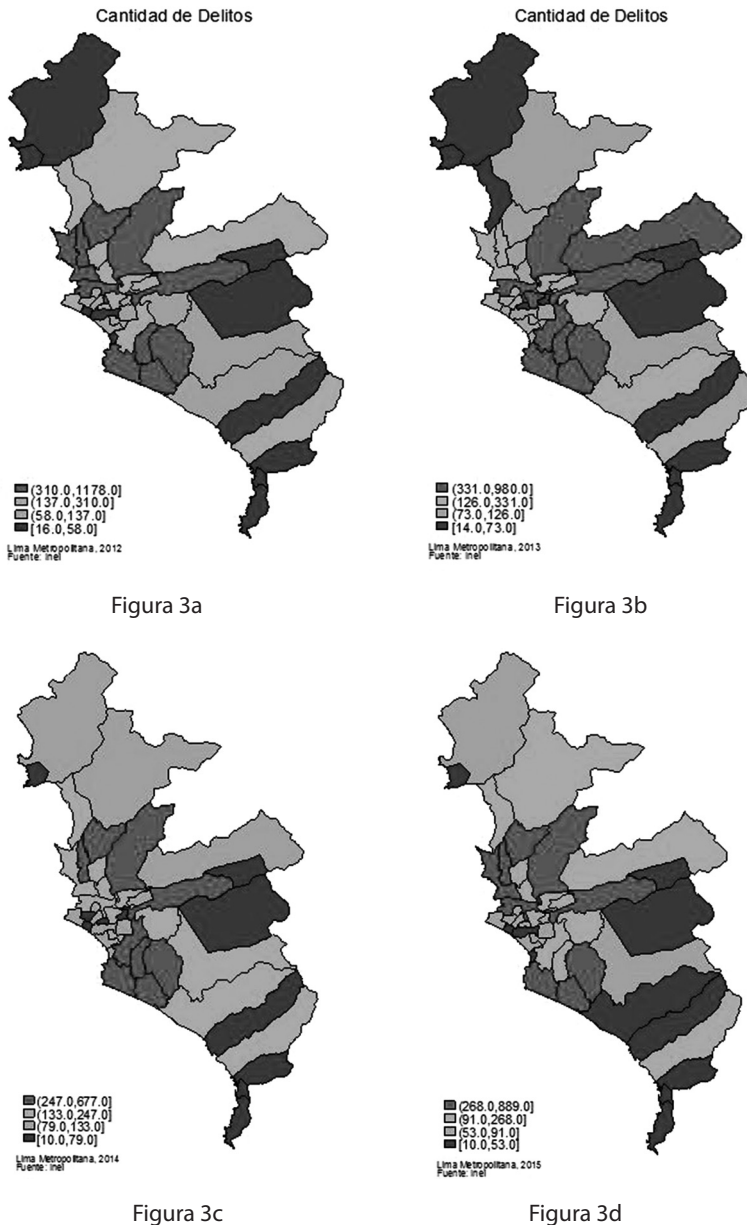
```
ycord(y_c) deviation( delitos2013) fcolor(red) size(*0.6)) legend(size(medium) position(5)) ///
```

```
note("Lima Metropolitana, 2015" "Fuente: Inei")
```

### Evolución de delito por año:

A continuación se presentan los mapas de coropleto aplicados a los datos de INEI utilizando la información del último año disponible:

Figura 3. Mapa de cuantiles



El mapa de cuantiles se genera utilizando el subcomando **clmethod(q)**. Además, puede modificarse el número de clases usando **cln(#)**, así como los colores, la posición de la leyenda, títulos, notas, etc. La primera línea previa al comando del gráfico, **format**, permite controlar el formato numérico de la leyenda.



La evolución va desde la Figura 3a hasta la Figura 3d, esta evolución muestra cómo los distritos se han contagiado en términos de delincuencia (como cantidad de delitos). Esto será analizado no solo gráficamente sino con un modelo de econometría espacial que estime el grado de correlación con el cual se mediría el efecto contagio.

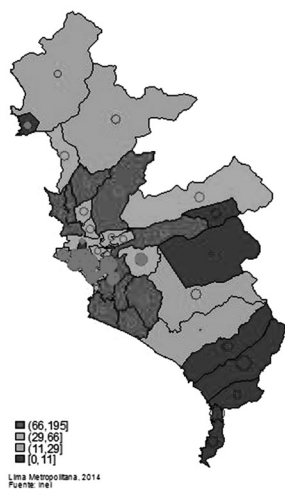


Figura 4a  
Número de delitos vs Ingreso promedio

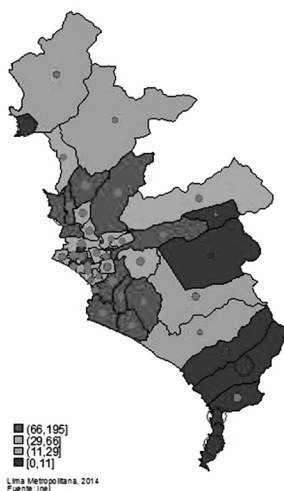


Figura 4b  
Número de delitos vs Tasa de desempleo

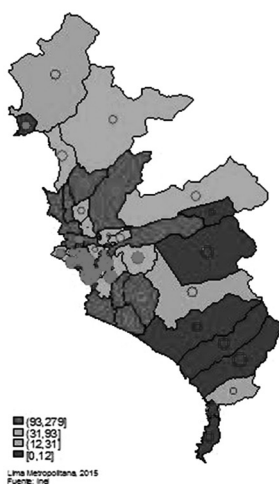


Figura 4c  
Número de delitos vs Ingreso promedio

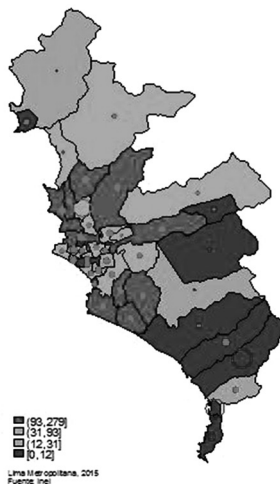


Figura 4d  
Número de delitos vs Tasa de desempleo

Los círculos huecos representan las desviaciones por debajo de la media del número de delitos y los círculos rellenos representan desviaciones por encima de la media. A mayor tamaño del círculo, más alejado se encuentra el valor de la región respecto a la media de Lima Metropolitana. Es fácil observar que los círculos huecos y de gran tamaño, que identifican a los números de delitos, se ubican en regiones de alto desempleo. Por el contrario, los círculos que identifican bajos números de delitos se ubican en áreas de mayor desempleo.

## Creación de la matriz $W$ y contraste de autocorrelación espacial

La subsección previa ha brindado los primeros indicios de dependencia o autocorrelación espacial, sin embargo, no hay una cuantificación de esta autocorrelación y no se conoce si la misma es relevante. Es entonces necesario el uso de herramientas inferenciales para obtener una cuantificación de la misma. En este punto surge el problema de la matriz de pesos espaciales.

En Stata, las matrices espaciales pueden generarse mediante diferentes comandos. Uno de ellos es **spmat** (Drukker et al., 2013b) que permite crear, importar, manipular y guardar matrices  $W$ . Las matrices son almacenadas como objetos **spmat**, estructura requerida para el uso de otros comandos como **spreg** (Drukker et al., 2013d) y **spivreg** (Drukker et al., 2013c).

El comando **spatwmat** (Pisati, 2001) está integrado a las herramientas de análisis exploratorio que incluye los contrastes de dependencia espacial global y, además, al análisis de dependencia local. Si la matriz ha sido generada por **spmat** no podrá ser usada directamente para los comandos del análisis exploratorio de Pisati.

Una tercera opción para matrices espaciales es **spwmatrix** de Jeanty (2014). Este comando permite generar matrices similares a **spmat**, pero además genera matrices de  $k$  vecinos más cercanos y matrices con criterios socioeconómicos.

El comando **spmat** ofrece varias ventajas. Entre ellas, permite la visualización de las matrices, así como alternativas avanzadas de especificación.

Veamos cómo trabajan estos comandos bajo Stata.

- `spmat contiguity Wcontig using "coord.dta", id(id)`
- `spwmatrix gecon SHAPE LENG SHAPE_LE_1, wn(W5st) knn(5) row`

Se necesita la matriz  $W$  como objeto SPMAT: Para ello se genera  $5nn$  binaria y se exporta a `txt`

- `spwmatrix gecon y_c x_c, wn(W5bin) knn(5) xport(W5bin,txt)`

Se lee el archivo y se adapta el formato para SPMAT

- `insheet using "W5bin.txt", delim(" ") clear`
- `drop in 1`
- `rename v1 id`
- `save "W5bin.dta", replace`

Se genera el objeto SPMAT: W5 estandarizado por filas

- `spmat dta W5_st v*, id(id) norm(row)`
- `spmat summarize W5_st, links`
- `spmat graph W5_st`

Una vez definida la matriz de contactos, pueden utilizarse diferentes estadísticos univariantes que permiten detectar autocorrelación espacial. Los mismos pueden clasificarse como medidas de dependencia global o local. Las medidas globales utilizan la información completa del conjunto de datos con el fin de obtener un valor promedio para todo el espacio geográfico. Al resumir en un único valor toda la información, no es posible detectar la variabilidad de la dependencia ni la localización de estos patrones. Por su parte, las medidas locales examinan la autocorrelación espacial en un subconjunto de datos. En este escrito nos focalizaremos en medidas globales y se dejará para próximos avances las medidas locales.

## Estadísticos globales

Posiblemente, el contraste ampliamente utilizado sea el estadístico  $I$  de Morán (1950):

$$I = \frac{n}{S_0} \frac{\sum_i \sum_j (y_i - \bar{y}) w_{ji} (y_j - \bar{y})}{\sum_{i=1}^N (y_i - \bar{y})^2}$$

Cuando el contraste  $I$  de Moran toma un valor positivo existe autocorrelación positiva implicando que los valores de cada observación y sus vecinos se asemejan. Si el  $I$  asume un valor negativo entonces esto implica autocorrelación negativa tal que el valor de los vecinos son altos cuando la observación tiene un valor bajo y si es alto entonces sus vecinos asumen valores bajos.

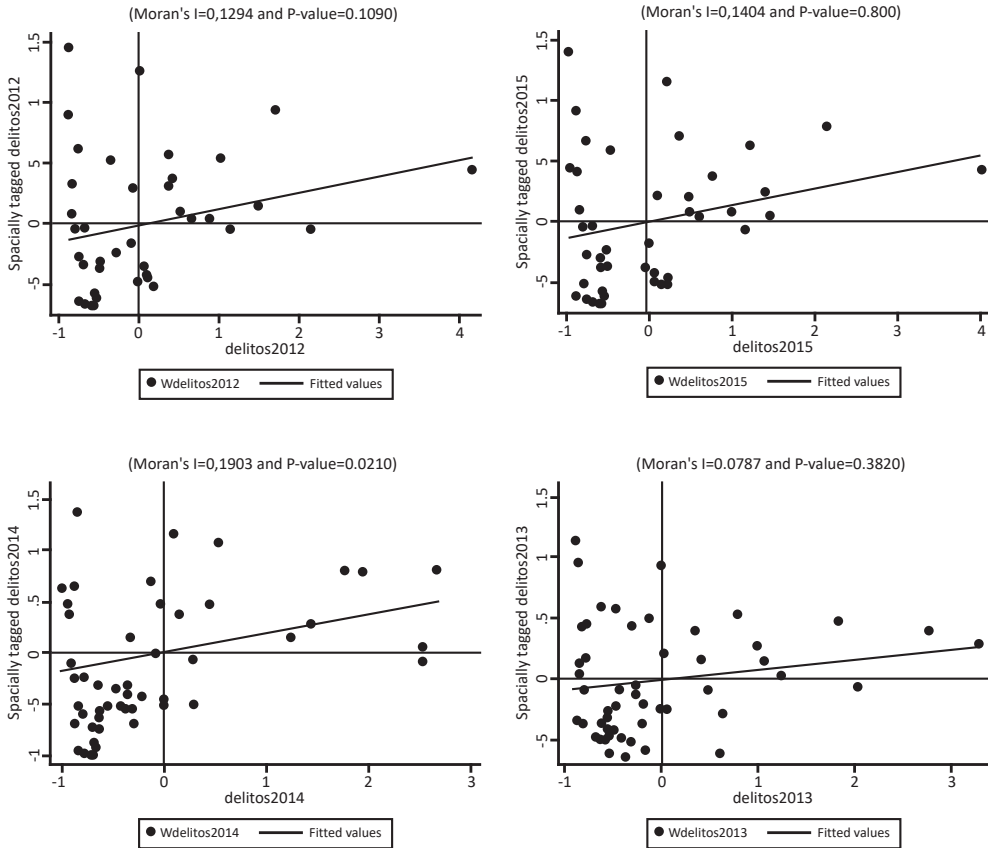
## Diagrama de dispersión de Moran

El diagrama de Morán divide al gráfico en cuatro cuadrantes respecto al valor promedio: valores altos de delictividad con valores altos de sus vecinos (Alto-Alto), valores bajos de delictividad con valores bajos de sus vecinos (Bajo-Bajo), valores altos de delictividad con valores bajos de sus vecinos (Alto-Bajo), y valores bajos del delictividad con valores altos de los vecinos (Bajo-Alto).

Los siguientes diagramas presentan un análisis de la dispersión según el índice de Moran

Gráfico 6

Diagrama de dispersión de Morán



Conclusiones

El efecto contagio de la delincuencia se ve reflejado en el índice de correlación (Índice de Moran), un alto índice de correlación espacial entre distritos como San Juan de Lurigancho y distritos aledaños muestra que la delincuencia se expande de manera territorial hacia ese distrito, esto prueba la idea de un efecto del territorio sobre los actos delictivos, de manera simple se puede concluir que distritos con altos índices delincuenciales tendrían influencia en el incremento de índices delincuenciales en distritos limítrofes.

El modelo comprueba espacialmente la relación directa que hay entre pobreza y delincuencia, muestra relaciones de dependencia espacial entre estas variables, el modelo muestra que distritos pobres (bajos ingresos) tienen más probabilidad de tener como vecinos limítrofes a distritos con bajos ingresos y de manera análoga vecinos con

alta actividad delictiva. De acuerdo a los indicadores de correlación simple se nota que existe una fuerte correlación entre el ingreso promedio por distrito y el número de actos delictivos, así mismo entre el número de actos delictivos y el nivel de desempleo. De esta manera, el nivel de desempleo, visto de una manera espacial, es considerada como una posible variable causante de actos delictivos y de correlación espacial entre regiones.

De acuerdo al análisis espacial considerando las diferencias entre áreas urbanas y rurales, la violencia afecta mucho más las áreas urbanas que en las rurales, principalmente las ciudades grandes que a las pequeñas, como por ejemplo en Lima se notó que la presencia de mayor delictividad se encuentra principalmente en los distritos de San Juan de Lurigancho, Ate Vitarte y el Callao.

Los actos delictivos estudiados así como la correlación espacial nos llevan a pensar que la tendencia es positiva y no a la inversa, es decir que el bienestar de la población (medida por ejemplo como la percepción de la inseguridad entre otras variables claves) seguirá en decadencia por algunos años más, a menos que se tomen medidas correctivas. Los indicadores de homicidios, robos y otros delitos nos muestran el impacto de los crímenes sobre la salud de las personas en las lesiones o en la muerte prematura.

## **Bibliografía:**

Chasco, C. *Econometría espacial aplicada a la predicción-extrapolación de datos microterritoriales. Ganadora de la Quinta Edición de Premios sobre Estudios de Economía Regional. Recuperado:* Link: <http://www.madrid.org/bvirtual/BVCM005618.pdf>

Formisano, M. *Econometría espacial: características de la violencia homicida en Bogotá.* Recuperado: Link: <https://core.ac.uk/download/pdf/6325280.pdf>

Jacobs, J. (2011). *Muerte y vida de las grandes ciudades.* Madrid: Capitán Swing Libros.

Peru 21. Población ocupada en Lima Metropolitana creció 0.6%, anunció el INEI- 15 de diciembre del 2016-. Link: <http://peru21.pe/economia/inei-poblacion-ocupada-lima-metropolitana-crecio-06-2265050>

## **Referencias bibliográficas**

Barón, L., Sandoval, E. y Martínez, D. (2008). *Una revisión al estudio de delincuencia y criminalidad.* NUEVA GRANADA. Recuperado el junio de 2008

Carlos J. Vilalta y Perdomo. (2003). Una aplicación del análisis espacial al estudio de las diferencias regionales del ingreso en México. *Economía, Sociedad y Territorio, IV(14)*. Obtenido de <https://est.cmq.edu.mx/index.php/est/article/view/429/843>

- Corman, H. & Mocan, N. (2000). En *A Time-Series Analysis of Crime, Deterrence, and Drug Abuse in New York City* (Vol. 90, págs. 584-604). The American Economic Review.
- David Martin. (2002). Patrones espaciales en robo residencial. En D. Martin, *Evaluar el efecto del capital social del vecindario* (págs. 132-146). Journal of Contemporary Criminal Justice . Obtenido de <http://journals.sagepub.com/doi/pdf/10.1177/1043986202018002002>
- Fajnzylber, Pablo, Daniel Lederman, Norman Loayza. (2002). What causes violent crime? En *European Economic Review* (E. autor, Trad., Vol. 46, págs. 1323-1357).
- Flores C., C. (2008). Lecciones de Criminología. Caracas-Venezuela: Vadell Hermanos Editores. Obtenido de <https://temasdederecho.wordpress.com/2012/06/04/la-criminologia-como-ciencia/>
- Formisano, M. (SEPTIEMBRE de 2002). Econometría espacial: Características de la violencia Homicida en Bogotá. En Cede (Ed.). Obtenido de <https://core.ac.uk/download/pdf/6325280.pdf>
- Herrera, M. (2015). *Econometría especial usando Stata. Breve guía aplicada para datos de corte transversal* (Vol. 13). Argentina: ielde - Facultad de Ciencias Económicas, Jurídicas y Sociales - UNSA.
- Llorente, Escobedo, Echeandía y Rubio (2000). Caracterización de la violencia homicida en dos localidades de Bogotá. En *Caracterización de la Violencia Homicida en Bogotá* (Vol. III). CEDE - Universidad de los Andes.
- tituto de Estudios Peruanos. Obtenido de <http://www.iep.org.pe/titulos4.php>
- Sánchez, Fabio; Díaz, Ana María; Formisano, Michel. (2003). En *Conflicto, violencia y actividad criminal en Colombia: Un análisis espacial* (págs. 1-60). CEDE. Obtenido de [https://economia.uniandes.edu.co/components/com\\_booklibrary/ebooks/D2003-05.pdf](https://economia.uniandes.edu.co/components/com_booklibrary/ebooks/D2003-05.pdf)
- Tudela, G. N. (5 de Mayo de 2011). Análisis de las actividades delictivas en Bolivia, desde la perspectiva económica. La Paz, Bolivia.
- Tudela, G. N. (Enero-Junio de 2012). Análisis de las actividades delictivas en Bolivia desde la perspectiva económica. *CLIO América*, VI(11). doi: <http://dx.doi.org/10.21676/23897848.424>
- Zárate, P., Aragón, J. y Morel, J. (2013). Inseguridad, Estado y desigualdad en el Perú y en América Latina. En *Un estado de la cuestión*. Lima: Instituto de Estudios Peruanos, p. 14.

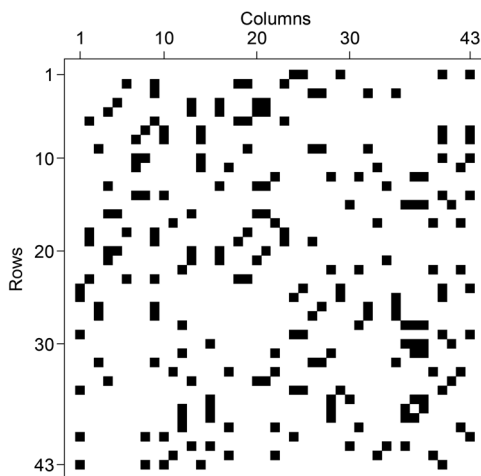
## ANEXOS

### Anexo 1. Reporte estadístico de las cantidad de delitos en Lima Metropolitana durante el periodo 2012-2015

Tabla  
sum delitos2012 delitos2013 delitos2014 delitos2015

Variable	Obs	Mean	Sed. Dev.	Min	Max
delitos2012	43	220.4419	230.4224	16	1178
delitos2013	43	221.6279	229.5054	14	980
delitos2014	43	195.814	180.7788	10	677
delitos2015	43	178.6977	177.3099	10	889

### Anexo 2. Matriz de contiguidad



# La educación y la probabilidad de ser pobre en el Perú

Eduardo Quiroz Vera

## Resumen

El principal objetivo de esta investigación es analizar los factores educacionales, demográficos, geográficos, mercado laboral, vivienda, ingreso y patrimonio relacionados con el jefe de hogar y la pobreza a través de un modelo logit, tratando de explicar la probabilidad de ser pobre monetario. Para esto se analizó los datos recogidos por la Encuesta Nacional de Hogares del 2016 realizada por el INEI. A través del primer modelo logit se encuentra que la educación –por sus efectos sobre la productividad y la generación de ingresos– se constituye en un instrumento clave en la política de superación de la pobreza, puesto que si mayor es el nivel alcanzado en sus estudios, mayores son las reducciones en la probabilidad de ser pobre; en especial se encuentra que concluir la universidad significa reducir la probabilidad de ser pobre en 14,1 puntos porcentuales, con respecto a un individuo que concluye secundaria, mientras que concluir una carrera técnica significa reducir la probabilidad de ser pobre en 9 puntos porcentuales con respecto a un individuo que concluye secundaria. En el segundo modelo logit queda demostrada la importancia de la educación en la probabilidad de ser pobre, pero de que por sí sola no podrá exhibir retornos positivos si el diseño de las políticas públicas no son eficientes y no se consideran algunos aspectos como son los aspectos demográficos, laborales, patrimoniales, geográficos y de vivienda, que son los que también explican la probabilidad de ser pobre.

**Palabras clave:** Educación, pobreza monetaria, Logit



## Introducción

La relación entre educación y pobreza se muestra en diferentes estudios empíricos efectuados por académicos y entidades multilaterales, pero desde el punto de vista teórico no existe un modelo que nos explique cómo se da la relación funcional entre ellas, y esto es importante para el desarrollo de una adecuada política social por parte de los gobiernos de turno.

Como indica Aguado (2007), si el grado de pobreza depende decrecientemente del nivel educativo, entonces se requiere de una política que lo incremente en los sectores más necesitados de la sociedad. Por el contrario, si el nivel educativo está supeditado de manera decreciente al de pobreza, entonces se precisa una política social que atenúe dicha situación. Finalmente, si la relación de causalidad implica retroalimentación entre los niveles de pobreza y de educación —es decir, si se es pobre por detentar bajos grados de educación y viceversa— el diseño de política social debería considerar de forma conjunta la dinámica educación-pobreza.

En efecto, de acuerdo con las últimas mediciones de la pobreza en América Latina (CEPAL, 2016), las tasas de pobreza e indigencia en América Latina se incrementaron el año pasado y alcanzaron a 175 millones de personas. De 2014 a 2015 la tasa regional de pobreza habría aumentado a 29,2% a nivel regional y la tasa de indigencia a 12,4%.

El solo crecimiento económico no soluciona el problema de la pobreza y la educación representa la clave para resolverlo de manera exitosa. Por consiguiente, resulta imperioso garantizar la inclusión de los más pobres en el sistema educativo, ya que a mayor nivel de educación son más altas las probabilidades de superar la pobreza. Según se ha estimado, un individuo reduce en un 6 % la probabilidad de ser pobre por cada año de educación.

Según CEPAL la inversión en capital educativo es un factor esencial para la reducción de la pobreza y la desigualdad, sobre todo por su capacidad de contribuir a la movilidad social y a la ruptura de la transmisión intergeneracional de la privación. CEPAL calcula que se necesita por lo menos 12 años de escolarización para conseguir escapar de la pobreza. Este es un dato revelador que no solo afecta a los países en desarrollo.

Especialistas del Banco Mundial afirman que “la educación contribuye al crecimiento económico, pero no lo genera por sí sola” (Fuente), y expertos de distintas latitudes coinciden en el papel preponderante de la educación para el desarrollo social y económico de una nación. La educación es considerada como la inversión social con las más altas tasas de retorno, tanto para la sociedad como para los individuos.

Hoy es evidente el modo en que se encuentra dividido el mundo por el grado de desarrollo tecnológico y científico de las naciones; por esa razón, para el país resulta políticamente imprescindible prestar atención a la educación como factor para lograr un mayor desarrollo económico y social.

El principal objetivo del presente trabajo es analizar los factores educacionales y demográficos relacionados con el jefe de hogar y la pobreza para tratar de explicar

la probabilidad de ser pobre. La hipótesis general planteada es si ser pobre o no está determinado por un conjunto de características estructurales del hogar vinculadas con dimensiones geográficas, demográficas, mercado laboral, educación, vivienda, ingresos y patrimonio.

El documento está organizado de la siguiente manera. El capítulo 2 discute y analiza trabajos que anteriormente han estudiado el impacto del nivel educativo alcanzado, ya sea a nivel de hogares o de individuos, en la probabilidad de ser pobre. El capítulo 3 resume y describe el marco teórico conceptual de referencia. El capítulo 4 discute la estrategia empírica llevada a cabo y explica por qué la metodología elegida es la adecuada, así como las consideraciones metodológicas con respecto al cálculo de la pobreza en el Perú considerando el objetivo y la hipótesis del estudio, además de la fuente de datos utilizada para las estimaciones. El capítulo 5 analiza los principales hallazgos de la investigación. Por último, el documento concluye con la discusión de los resultados empíricos y las conclusiones.

## **Breve revisión de la literatura**

La investigación guarda relación con los estudios que tienen como objetivo analizar la relación entre educación y pobreza. Considera en primer lugar, de una manera sencilla e intuitiva, aunque con limitaciones, el análisis de Crosstabs que permite identificar el impacto del nivel educacional del jefe de hogar sobre el nivel de pobreza del hogar; y en segundo lugar, considera la estimación econométrica de un modelo de probabilidad que permita medir la importancia relativa de la educación (tipos y niveles), y al mismo tiempo controlado por otras variables en la determinación de la pobreza.

Yamada y Castro (2008) evalúan el rol que tiene el grado de instrucción del individuo en Guatemala como determinante en su situación de pobreza, a través de la Encuesta Nacional de Condiciones de Vida (Encovi) para el año 2006. Los efectos impacto referidos a cada variable en educación son negativos y, más importante aún, crecientes en valor absoluto. En particular, un adulto sin educación tiene un 54 % de probabilidad de ser pobre, a diferencia de alguien con educación primaria completa, cuya probabilidad es de 23 %. Además, si el individuo posee educación secundaria completa, su probabilidad de ser pobre se reduce a 6 % y, si ha culminado estudios universitarios, esta probabilidad es de tan solo 1 %. Esto significa que el impacto marginal de concluir la educación básica (48 puntos porcentuales) con respecto a la educación superior (5 puntos porcentuales) es mayor, lo cual debería confirmar la necesidad de expandir la oferta de educación básica pública como mecanismo para igualar las oportunidades de generación de ingresos en Guatemala. No obstante, también encontraron que la probabilidad de ser pobre se ve afectada por características propias del individuo como la raza, la cual tiene un impacto positivo significativo (alrededor de 18 puntos porcentuales).

Sánchez (2015) evalúa el impacto de la educación sobre la pobreza en Costa Rica, por medio de la metodología Propensity Score Matching y los datos de la Encuesta Nacional de Hogares 2013. Se halla un efecto positivo de la educación en la reducción de la pobreza y se concluye que cuando una persona promedio cuenta con educación secundaria completa, reduce la probabilidad de encontrarse en pobreza en 5,7 puntos porcentuales, lo que representa una disminución en las tasas de pobreza de quienes han finalizado secundaria de 21,1 %. Por otra parte, la posibilidad de realizar estudios de postsecundaria disminuye la probabilidad de pobreza en 8,4 puntos porcentuales, y es equivalente a una disminución de 36,8 % en las tasas de pobreza de los individuos que al menos han terminado la educación secundaria.

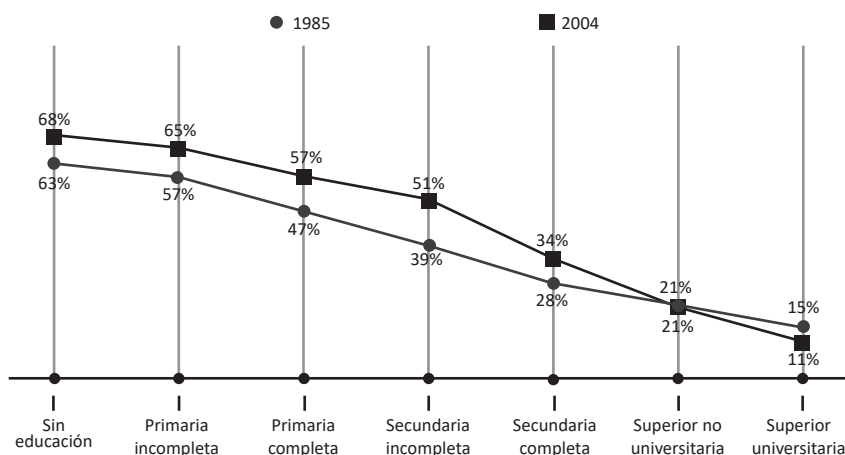
Bernal (2015) analiza el impacto de la educación sobre la pobreza y la desigualdad en Bolivia, y encuentra que a mayor nivel educativo aprobado, menor es la probabilidad de ser pobre, pues la contribución marginal de cada año de estudio reduce la probabilidad de ser pobre en 1,84%. El efecto marginal de tener una profesión universitaria reduce la probabilidad de ser pobre en 14,09 % y si se tiene un postgrado la probabilidad de ser pobre se reduce en 33,48 %. Asimismo, se comprueba la incidencia de la variable experiencia, hace significativo el aporte de esta variable (cada año adicional de experiencia, *ceteris paribus*, reduce la probabilidad de ser pobre en 2,69 %). El área de residencia y el idioma que habla es de importancia, si habla algún idioma originario aumenta la probabilidad de ser pobre en 3,96 %. El hecho de vivir en zonas rurales aumenta la probabilidad de ser pobre en 8%. Si el jefe del hogar es casado disminuye la probabilidad de ser pobre en 0,43 %. Por otro lado, concluye también que pasar de un nivel educativo a otro aumenta el ingreso en 1,28 %, vivir en el área rural disminuye el ingreso en 63 %; asimismo si habla un idioma originario disminuye el ingreso 10,56 % y si el jefe de hogar es analfabeto disminuye el ingreso en 2 %.

Al estudiar el caso peruano, Nelson Shack<sup>1</sup> analiza la información recogida por la Encuesta de Niveles de Vida (ENNIV) realizado por el Instituto Cuánto en 1997, a través de un modelo probit de máximoverosimilitud y concluye que el nivel de educación alcanzado por el jefe de hogar es importante para explicar la probabilidad de ser pobre y que su relevancia depende de la tasa de retorno con la que el mercado laboral la remunera.

Yamada (2005) encuentra evidencia mixta acerca de la rentabilidad de la educación superior en el mercado laboral peruano. Gracias a la convexificación de los retornos, la educación universitaria (tanto privada como pública) ofrece actualmente tasas de retorno privado y social en términos reales comparables con otras alternativas de inversión financiera y económica disponibles en el país. Por el contrario, la educación superior no universitaria otorga tasas de retorno privado y social poco atractivas, y hasta negativas, por lo que se requiere una profunda evaluación y reforma de este nivel educativo. El estudio encuentra también diferencias en los retornos a la educación de acuerdo al género, tipo de empleo y ubicación geográfica de los trabajadores.

Yamada y Castro (2007) indican que dada la estrecha relación educación-pobreza, las probabilidades de ser pobre aumentan para quienes no tienen buena educación. El 34 % de personas sin educación es pobre extremo, mientras que solo el 9 % de quienes tienen secundaria lo es. El problema se agrava para los que reciben educación pública. Según Yamada, la brecha de retornos entre la educación privada y pública aumentó de 1,1 puntos porcentuales (pp) en 1985-1986 a 8.6 pp en 2004, la educación superior no universitaria genera tasas de retorno privado y social cercanas a cero. En 1985, una persona con secundaria incompleta tenía una probabilidad de 39 % de ser pobre, hoy su probabilidad es de 51 % (Gráfico 1).

Gráfico 1  
Probabilidad de ser pobre según el nivel educativo



Fuente y elaboración: Yamada y Castro (2007).

Morón, Castro y Sanborn (2009) estudian la probabilidad de ser pobre monetario en Perú según el nivel educativo alcanzado, pero diferencian por educación pública y privada. Los resultados fueron obtenidos utilizando los datos de la Encuesta Nacional de Hogares (Enaho) del año 2006 y mediante una regresión logística, en la cual solo se consideró la población adulta. Encontraron que los individuos que concluyeron la educación básica en centros

privados son menos propensos a ser pobres a diferencia de aquellos que estudiaron en centros estatales. Por ejemplo, un individuo que terminó la primaria en un centro educativo estatal tiene una probabilidad de 43 % de ser pobre, a diferencia de alguien que terminó primaria en un centro privado, cuya probabilidad es de 15 %, es decir, 28 puntos porcentuales menos. Una situación similar se observa en educación secundaria, pues alguien que terminó la secundaria en un centro estatal tiene una probabilidad mayor de ser pobre que aquel que terminó sus estudios en un centro

privado. No obstante, no se observan diferencias significativas con respecto a la educación superior, lo cual implica que alcanzar dicho nivel educativo es una ruta segura de escape de la pobreza.

Salazar, Quispe y Choque (2015) realizan un análisis de la probabilidad de ser pobre monetario para una persona en función al nivel educativo alcanzado. Su análisis parte de la Encuesta Nacional de Hogares y determinan que concluir la universidad significa reducir la probabilidad de ser pobre en 12 puntos porcentuales, con respecto a un individuo que concluye secundaria; y 38 puntos porcentuales con respecto a un individuo sin educación.

Asimismo, se debe remarcar el carácter regresivo del acceso a la educación superior, pues significa una promesa de un beneficio económico pero especialmente para aquellos hogares que ya se encuentran en una situación económica estable.

## **Teoría para este estudio**

### **Conceptos y enfoques**

La pobreza es una condición en la cual una o más personas tienen un nivel de bienestar inferior al mínimo socialmente aceptado. En una primera aproximación, la pobreza se asocia a la incapacidad de las personas para satisfacer sus necesidades básicas de alimentación. Luego se considera un concepto más amplio que incluye la salud, las condiciones de vivienda, educación, empleo, ingresos, gastos y aspectos más extensos como la identidad, los derechos humanos, la participación popular, entre otros (INEI, 2017).

En general, todas las definiciones apuntan a la situación en que se encuentran las personas que no disponen de los medios (“de producción”, activos físicos, intelectuales, sociales, culturales, financieros y demás) que permitan su desarrollo personal y reflejen el estilo de vida de la formación social en su conjunto.

Existen tres grandes enfoques para medir la pobreza. El primero es el enfoque de la pobreza absoluta, que toma en cuenta el costo de una canasta mínima esencial de bienes y servicios, y considera como pobres a todos aquellos cuyo consumo o ingreso se halla por debajo de este valor.

El enfoque de pobreza relativa considera al grupo de personas cuyo ingreso se encuentra por debajo de un determinado nivel. Por ejemplo, en algunos países se considera pobres a todos aquellos que tienen remuneraciones inferiores a la mitad del ingreso promedio (criterio aplicado en sociedades que han logrado erradicar la pobreza absoluta).

El enfoque de la exclusión social, de absoluta vigencia en Europa, presta atención a las personas que no pueden acceder a determinados servicios como por ejemplo el empleo, la educación superior, la vivienda propia, el empleo y otros.

Con respecto a los métodos de medición, solo nos enfocamos en el método de línea de pobreza, el cual centra su atención en la dimensión económica de la pobreza

y utiliza el ingreso o el gasto como medidas del bienestar. Al determinar los niveles de pobreza, se compara el valor per cápita de ingreso o gasto en el hogar con el valor de una canasta mínima denominada línea de pobreza.

Cuando se utiliza el método de línea de pobreza por el consumo, se incorpora el valor de todos los bienes y servicios que consume el hogar, indistintamente de la forma de adquisición o consecución. La utilización del gasto de consumo tiene la ventaja de que es el mejor indicador para medir el bienestar porque se refiere a lo que realmente consume un hogar y no a lo que potencialmente puede consumir cuando se mide por el ingreso. Otro aspecto favorable es que el consumo es una variable más estable que el ingreso, lo que permite una mejor condición de la tendencia del nivel de pobreza.

Así como existen enfoques y métodos para medir la pobreza, existen definiciones que nos permitirán centrar mejor la percepción del nivel de pobreza que puede tener una persona o un hogar en particular.

### **Pobreza absoluta**

Comprende a las personas cuyos hogares tienen ingresos o consumo per cápita inferiores al costo de una canasta total de bienes y servicios mínimos esenciales.

Línea de Pobreza Absoluta (LPA): Es el costo de una canasta mínima de Bienes (incluido los alimentos) y servicios.

### **Pobreza absoluta**

Consumo Hogar < Costo Canasta Básica Consumo (LPA)

### **Pobreza extrema**

Comprende a las personas cuyos hogares tienen ingresos o consumos per cápita inferiores al valor de una canasta mínima de alimentos.

Línea de Pobreza Extrema (LPE): Es el costo de una canasta mínima de alimentos.

### **Pobreza extrema**

Consumo Hogar < Costo Canasta Básica Alimenticia (LPE)

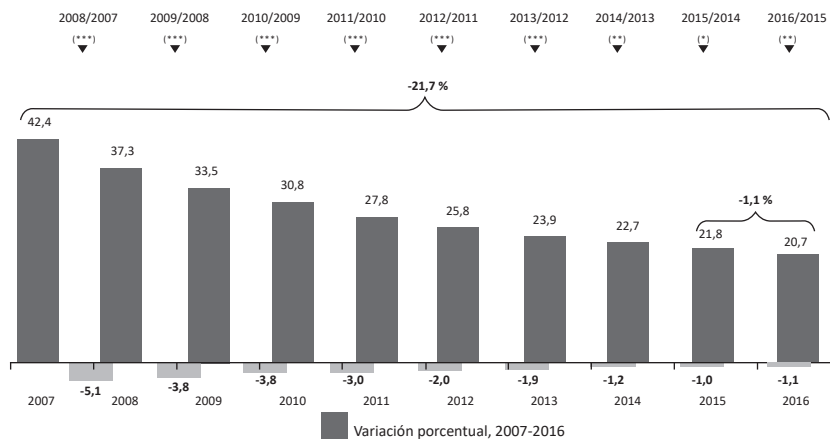
Según Sen (1992) la medición de pobreza requiere realizar dos ejercicios distintos pero interrelacionados: la identificación de los pobres por un lado y la agregación por otro. Este último es el usado para obtener indicadores resumen del nivel (incidencia) de la pobreza. Para el caso de esta investigación solo realizaremos en alguna medida el primer ejercicio.

## **La pobreza en el Perú**

A pesar de que la pobreza se redujo significativamente durante los últimos años (Gráfico 2), existen aún fuertes críticas a la política social en el Perú (INEI, 2017).

Gráfico 2

**PERÚ: EVOLUCIÓN DE LA INCIDENCIA DE LA POBREZA TOTAL, 2007-2016**  
**(Porcentaje respecto del total de población)**



Año	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016
Límite Inf.	41,1	36,0	32,2	29,5	26,7	24,7	23,0	21,8	20,9	19,9
Límite Sup.	43,8	38,6	34,8	32,0	28,9	26,9	24,8	23,6	22,7	21,6

\* Diferencia significativa ( $p < 0,10$ )  
 \*\* Diferencia altamente significativa ( $p < 0,05$ )  
 \*\*\* Diferencia muy altamente significativa ( $p < 0,01$ )  
 Fuente: INEI – Encuesta Nacional de Hogares, 2007-2016

El presupuesto público destinado al gasto social se ha incrementado fuertemente durante los últimos años, se han creado nuevos programas sociales y otros se han rediseñado; sin embargo, la debilidad de la organización institucional de la lucha contra la pobreza y la inadecuada gestión pública no han permitido alcanzar reducciones mayores en los niveles de pobreza. Asimismo, la política social en el Perú no le ha puesto el énfasis adecuado al combate de los determinantes de largo plazo de la pobreza y a la búsqueda de la disminución de la vulnerabilidad de los pobres.

Según Chacaltana (2006) resulta importante entender que la pobreza es heterogénea (todos los pobres son distintos entre sí) y que es dinámica (muchos individuos entran y salen continuamente de esta situación). Además, según este autor, es importante reconocer que la pobreza tiene determinantes de corto plazo (pobreza transitoria) y de largo plazo (pobreza crónica). Los determinantes de corto plazo de la pobreza, por ejemplo, pueden estar asociados a *shocks* temporales como desastres naturales o periodos de desempleo transitorio (relacionados con el ciclo económico) que generan que el hogar caiga temporalmente por debajo de la línea de pobreza. Por otro lado, los determinantes de largo plazo de la pobreza estarían asociados, por ejemplo, a los efectos demográficos, al acceso a diferentes activos productivos, al acceso a servicios públicos (agua, desagüe, telecomunicaciones y energía eléctrica), al

nivel educativo del jefe de hogar, al acceso a centros de salud, acceso a financiamiento y a otros factores que incidan sobre la productividad, oportunidades y capacidades de los miembros del hogar.

Así, se debería esperar que la pobreza disminuya de manera significativa en la medida en que se combatan los determinantes de largo plazo de la pobreza, en lugar de los determinantes de corto plazo de la misma.

En el mes de marzo de 2012, el Instituto Nacional de Estadística e Informática (INEI) dio a conocer al público usuario las razones de la actualización metodológica en la medición de la pobreza monetaria. Esta actualización se dio dentro de la política del INEI de garantizar la mayor transparencia y calidad de la información que produce y difunde. Este proceso es producto de un trabajo iniciado en el año 2007 con la conformación de un Comité Asesor Especializado de Pobreza con el auspicio del Banco Mundial. El Comité Asesor está conformado por representantes de organismos internacionales, de organismos gubernamentales nacionales, de la comunidad académica y de centros de investigación, que vienen participando cada año en la medición de la pobreza.

## **Razones para la actualización de la metodología de medición de la pobreza monetaria**

Los indicadores de pobreza monetaria que difundió el INEI hasta el año 2010 tuvieron como año base la Encuesta Nacional de Hogares de 1997. En ese año se definió la línea de pobreza base, la cual se mantuvo constante año tras año. Desde entonces hasta el año 2010, la economía y la sociedad peruana experimentaron profundos cambios demográficos, sociales y económicos. Además, en ese año se dispuso de información actualizada proveniente del Censo de Población y Vivienda del año 2007 y de la Encuesta Nacional de Presupuestos Familiares 2008-2009, para incorporarse como insumo en la nueva medición de la pobreza. Por estas razones, a partir del 2007 y en los años subsiguientes, el INEI, conjuntamente con la Comisión Consultiva de la Pobreza, trabajaron en algunas mejoras metodológicas a fin de hacer a la medición de la pobreza más consistente y acorde con la realidad. Los cambios ocurridos han modificado de manera significativa varios de los parámetros que entran en la definición de los indicadores de pobreza. A continuación se presentan los cambios más relevantes para la medición de la pobreza monetaria en el Perú:

- A. Cambios en la estructura poblacional urbano/rural
- B. Cambios en los patrones de consumo
- C. Cambios en los requerimientos calóricos
- D. Cambios en la estimación del gasto de los hogares
- E. Nueva población de referencia



## Metodología

Se parte de la teoría del capital humano y la definición de la pobreza monetaria para determinar si un individuo es pobre o no, se debe comparar el gasto mensual por persona, la cual es una variable continua y observable, con la canasta básica de consumo, la cual está compuesta por productos que son consumidos con mayor frecuencia, que representan la gran parte del gasto de los hogares, y que son homogéneos a escala nacional, es decir, tienen la misma unidad de medida e igual calidad (INEI, 2013).

En la parte empírica, la manera robusta de cuantificar la importancia relativa de la educación y de otras variables en la determinación de la situación final de pobreza de una persona requiere de la estimación econométrica de un modelo de probabilidad. Si se estima una regresión donde, por ejemplo, las variables explicativas sean los niveles de educación, el sexo, la edad, etc., y la variable explicada una de carácter dicotómica que refleje el hecho de ser pobre (la variable toma valor 1) o no (toma el valor 0), de acuerdo con determinada forma funcional, lo que se obtendrá es justamente la probabilidad de ser pobre. Más aún, si se deriva parcialmente la variable explicada ( $P$ , la probabilidad de ser pobre) con respecto a alguna de las explicativas (por ejemplo, años de educación) se obtiene el impacto de un año más de educación sobre la probabilidad futura de ser pobre.

Para la elección del modelo se toma en cuenta primero el tipo de función de distribución acumulada que utiliza; segundo, si las respuestas de los individuos pueden ordenarse o no; y tercero, por las variables consideradas explicatorias. A partir de lo anterior, los modelos más usados son el Logit, que utiliza la función de distribución acumulada logística; el Probit, que usa la función de distribución acumulada normal; y el de valor externo, que aplica la función de distribución acumulada de Gompit.

Para este estudio se escogió un modelo Logit. Cabe indicar que la elección de uno u otro modelo es arbitraria y su diferencia es fundamentalmente operativa, pues el objetivo primordial es la obtención de la probabilidad de que un individuo se clasifique en una categoría, según el valor de las variables explicativas. En ese sentido, los coeficientes estimados no cuantifican directamente el incremento en la probabilidad, dado un incremento unitario en la correspondiente variable independiente. Sin embargo, el signo de los regresores sí indica la dirección del cambio y, naturalmente, cuanto más alta sea dicha pendiente, mayor será el impacto del cambio en la variable explicativa sobre el cambio en la probabilidad.

El análisis se efectuó en tres etapas.

En la **primera etapa** se hizo una aproximación sencilla y bastante intuitiva a las relaciones de las variables que nos brindan posibles explicaciones de la pobreza en un marco general y de los individuos jefe de hogar en este caso en particular. Entre ellas se encuentran aquellas de carácter sociodemográfico como el género, la edad, el estado civil, el nivel de educación, el alfabetismo (variables de naturaleza cualitativa), medidas en algunos casos en escala nominal-dicotómica como el sexo del jefe

de hogar (hombre-mujer) y en otros casos nominal-politómica como el estado civil (soltero-casado-conviviente, etc.). En otro tipo de escalas se encuentran aquellas que denotan un orden ascendente en sus categorías como lo es el nivel de educación (Sin Nivel- Educación Primaria, etc.) alcanzado por el jefe de hogar. Encontramos además variables de naturaleza cuantitativa como lo constituyen la edad y el indicador proxy de experiencia laboral, que no considera aquel caso en el que el jefe del hogar ha estudiado y trabajado, sino solo aquel tiempo en el cual el jefe de hogar solo trabaja.

Uno de los objetivos que perseguimos y que está implícito dentro de esta investigación es dejar en claro trabajamos con variables a diferentes escalas, y que de acuerdo con estas escalas se determinará la implementación y ejecución de las metodologías que servirán de herramientas para una explicación de la pobreza en el Perú en particular.

**En la segunda etapa** se usó el modelo logit para evaluar la probabilidad de que un individuo caiga en condiciones de pobreza como función de las características y los niveles educacionales del jefe de hogar.

La probabilidad de que un individuo sea pobre es definida por ( $P$ ) con la siguiente expresión (del tipo logístico):

$$P_{(pobre)} = 1 / (1 + e^{-Z})$$

Donde  $Z$  es la combinación lineal:

$$Z = B_0 + B_1 X_1 + B_2 X_2 + \dots + B_p X_p$$

$B_i$  son los coeficientes por estimar

$X_i$  son las variables independientes

$e$  es la base del logaritmo natural, aproximadamente 2,718

Si la probabilidad de que un individuo sea pobre es  $P$ , entonces la probabilidad de que no lo sea es  $(1 - P)$ , en consecuencia, la razón entre ambas probabilidades será igual a:

$$P / (1 - P) = (1 + e^Z) / (1 + e^{-Z}) = e^Z = \exp(Z)$$

De manera que llegamos a la siguiente ecuación logit o log-razón de probabilidad:

$$\ln[P / (1 - P)] = B_0 + B_1 X_1 + B_2 X_2 + \dots + B_p X_p$$

Entonces, tenemos que  $P_{(pobre)}$  como la probabilidad de caer en condición de pobreza, tomará valores de uno o cero según el individuo sea pobre o no lo sea.

En la **tercera y última etapa** se modeló una especificación ampliada a fin de recoger el impacto de variables laborales y demográficas del hogar sobre la probabilidad de ser pobre y evaluar la importancia de la educación en un marco de referencia más completo.

## **Información primaria por utilizar**

La base de datos empleada en esta investigación corresponde a la Encuesta Nacional de Hogares llevada a cabo por el Instituto Nacional de Estadística e Informática. Esta encuesta de hogares permite obtener información actualizada sobre las condiciones de vida, gastos e ingresos de los hogares y, fundamentalmente, conocer la evolución de los niveles de pobreza monetaria en el país. La cobertura es nacional, en el ámbito urbano y rural, y con desagregación departamental, y desde el año 2003 se realiza de forma continua. Además es importante resaltar que el tamaño de muestra a nivel nacional registró cambios en este periodo: de 2004 a 2006 fue de 21 919 viviendas, de 2007 a 2010 el tamaño fue de 22 640 viviendas. Finalmente, desde 2011 el tamaño de muestra es de 26 456 viviendas.

En el año 2012 se decidió actualizar las cifras de pobreza, pues hasta el año 2010 habían sido definidas de acuerdo con la información e indicadores recolectados en función a la encuesta del año 1997.

Finalmente, se considerará solo a los individuos de 25 años de edad o más, puesto que se estima que a esta edad una persona ya ha culminado el tránsito por el sistema educativo.

## **VARIABLES DE ESTUDIO**

La variable dependiente es el nivel de pobreza y puede tomar dos valores: pobre y no pobre. Las variables independientes consideran las siguientes dimensiones:

- Geográficas: Región y área.
- Demográficas: Sexo del Jefe(a) de Hogar (JH), número de personas en el hogar, número de menores de 14 años en el hogar
- Mercado Laboral: Actividad del JH (Ocupado, Desocupado, Inactivo), número de ocupados en el hogar.
- Educación: Nivel Educacional del JH y Años de Escolaridad del JH
- Vivienda: Calidad de la vivienda, acceso al agua potable, acceso al alcantarillado, acceso a la energía eléctrica, hacinamiento.
- Ingresos: número de perceptores de ingresos en el hogar
- Patrimonio: Tenencia de teléfono, computador, microondas, videograbador, lavadora, refrigerador, internet y celular.

## 5. Análisis de resultados: análisis correlacional y uso del modelo logit

A continuación se presentan los resultados de un análisis en tres etapas. La primera contiene una aproximación sencilla y bastante intuitiva a las relaciones entre la educación y la pobreza; a través de un Análisis de Tablas Cruzadas se evalúa la correlación entre el estado de la pobreza del jefe de hogar con la educación y algunas variables que guardan relación con la pobreza.

En la segunda etapa se procede a estimar econométricamente un modelo probabilístico que permita calcular las magnitudes en que dichas relaciones se presentan. Este análisis se efectúa mediante un Modelo Logit y considera únicamente las características educacionales individuales del jefe de hogar.

En la tercera etapa se procede a estimar un modelo ampliado que recoge el impacto de las variables demográficas, geográficas, vivienda, patrimonio y laborales del jefe de hogar sobre la probabilidad de ser pobre, con el propósito de analizar la importancia de la educación en un marco de referencia más completo.

### 5.1. Análisis de tablas cruzadas

En esta parte de la investigación se describen aquellas variables que nos brindan posibles explicaciones de la pobreza de los individuos jefe de hogar. En este caso en particular haremos uso de las variables nivel educativo, género, estado civil, área de residencia y dominio geográfico.

Una de las características de la pobreza es que esta se encuentra asociada a niveles bajos de educación alcanzados por el jefe de hogar. En la Tabla 1 se observa que 64,6 % y el 45,8 % de los jefes de hogar sin nivel educativo y nivel educativo primario, respectivamente, se encuentran en condición de pobreza.

Tabla 1

#### Perfil del jefe de hogar según nivel educativo y estado de pobreza

Nivel educativo	Estado de pobreza		Total
	Pobre	No pobre	
Sin nivel	64,6	35,4	100,0
Primaria	45,8	54,2	100,0
Secundaria	30,7	69,3	100,0
Superior no universitaria	17,1	82,9	100,0
Superior universitaria	6,5	93,5	100,0
Posgrado universitaria	0,03	99,97	100,0
Total	22,8	77,2	100,0

Fuente: ENAHO 2016-IV. Elaboración propia.

En el caso de los que tienen nivel educativo secundario, se observa que el estado de pobreza se encuentra repartido para sus dos categorías, con 30,7 % para el caso de jefes de hogar pobres y con 69,3 % para el caso de los no pobres. Una situación inversa se da en los niveles superiores, en los que el 82,9 % de jefes de hogar con nivel educativo superior no universitario se encuentra en estado de no pobreza. En una tendencia creciente mucho mayor y estado de no pobreza se encuentran aquellos jefes de hogar que tienen niveles como el universitario y el posgrado, con 93,5 % y 99,97 %, respectivamente.

En la Tabla 2 se observa que al aplicar la prueba Chi cuadrado para probar si el nivel educativo del jefe de hogar es independiente del estado de pobreza, se encontró que existe una relación significativa entre ambas variables ( $p < 0.05$ ); por lo tanto, a medida que el jefe de hogar adquiere un mayor nivel de educación, obtiene más capacidades para salir de la condición de pobreza; más aún si tiene niveles de educación superiores.

Por otro lado, al analizar los coeficientes de correlación (Tabla 2), todos ellos indican un fuerte grado de asociación entre el nivel educativo y el estado de pobreza (todos los coeficientes son estadísticamente significativos).

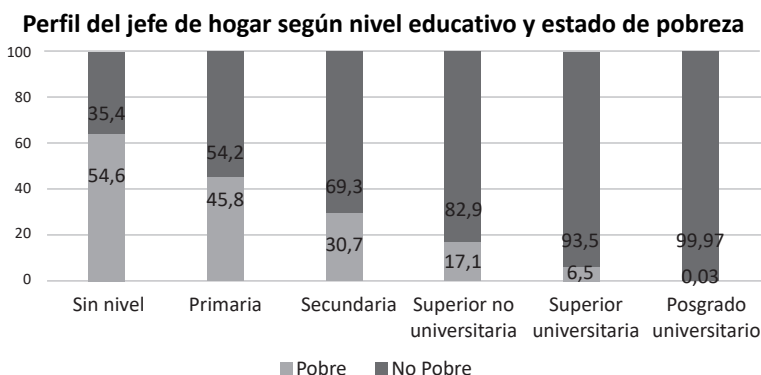
Tabla 2

**Valores y significancia de la prueba Chi cuadrado y medidas simétricas**

Prueba Chi cuadrado		Medidas simétricas		
		Coef. de correlación	Valor	Ap prox. Sig.
Pearson Chi cuadrado	69234,8	Kendall's tau-b	0,517	0,000
		Kendall's tau-c	0,483	0,000
		Gamma	0,521	0,000
Asymp. Sig. (2-sided)	0,000	Spearman Correlation	0,494	0,000

Fuente: Elaboración propia.

Gráfico 3



Se puede apreciar (Tabla 3) que dentro de la condición de pobreza, los jefes de hogar hombres constituyen el 90,9 % del estado civil conviviente, en igual sentido se encuentran los jefe de hogar casados que vienen a estar constituidos en un 92,5 % por los hombres. Para las otras categorías de estado civil cambia el sentido y se observa que en el caso de las mujeres para el estado civil viuda, divorciada, separada y soltera, constituyen el 77,4 %, 84,2 %, 61,2 % y 54,6 %, respectivamente.

De la tabla 4, se observa que la prueba Chi cuadrado es estadísticamente significativa, lo que nos permite inferir que existe relación significativa entre el estado civil y el sexo del jefe de hogar en condición de pobreza. También se observa que los coeficientes de correlación indican un fuerte grado de asociación entre el género y el estado civil en estado de pobreza, pues todos los coeficientes son estadísticamente significativos.

Tabla 3  
Perfil del jefe de hogar según género y estado civil en estado de pobreza

Estado civil	Sexo		Total
	Hombre	Mujer	
Conviviente	90,9	9,1	100,0
Casado	92,5	7,5	100,0
Viudo	22,6	77,4	100,0
Divorciado	15,8	84,2	100,0
Separado	31,8	61,2	100,0
Soltero	45,4	54,6	100,0
Total	68,9	31,1	100,0

Fuente: ENAHO 2016-IV. Elaboración propia

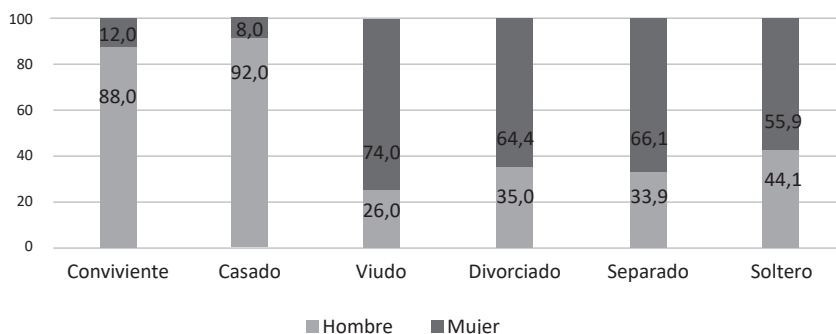
Tabla 4  
Valores y significancia de la prueba Chi Cuadrado y medidas simétricas

Prueba Chi cuadrado		Medidas simétricas		
Pearson Chi cuadrado	901847	Coef. de correlación	Valor	Ap prox. Sig.
		Phi	0,626	0,000
		V de Cramer	0,624	0,000
Asymp. Sig. (2-sided)	0,000	Coeficiente de contingencia	0,53	0,000

Fuente: Elaboración propia

Gráfico 4

**Perfil del jefe de hogar según género y estado civil en estado de no pobreza**



En el año 2016, podemos ver (Tabla 5) que en condición de pobreza los jefes de hogar del área urbana constituyen el 82,1 %, 82,4 % y 94,2 % para la región de la costa norte, costa centro y costa sur, respectivamente. En un nivel porcentual menor, pero igual de considerable, la sierra norte está constituida en un 52,8 % por jefes de hogar del área rural en condición de pobreza, mientras que en la sierra sur está constituida en un 52,8 % de jefes de hogar del área urbana en condición de pobreza. En la sierra centro no hay una diferencia entre los porcentajes de jefes de hogar según el área de residencia. Aunque en el área rural de la selva los jefes de hogar pobres son alrededor del 51,1 %, no es tan evidente la diferencia como en el resto de regiones, debido a que en el área urbana se encuentra un 48,9 % en estado de pobreza.

Tabla 5

**Perfil del jefe de hogar según área y dominio geográfico en estado de pobreza**

Dominio geográfico	Área		Total
	Urbana	Rural	
Costa norte	82,1	17,9	100,0
Costa centro	82,4	17,6	100,0
Costa sur	94,2	5,8	100,0
Sierra norte	47,2	52,8	100,0
Sierra centro	50,3	49,7	100,0
Sierra sur	52,8	47,2	100,0
Selva	51,1	48,9	100,0
Total	69,8	30,2	100,0

Fuente: ENAHO 2016-IV. Elaboración propia

Los datos muestran que la condición de pobreza afecta en mayor medida a aquellos que menos acceso tienen en cuanto a activos públicos de infraestructura, educación, salud, etc.; que son, justamente, aquellos jefes de hogar que se encuentran en la sierra y selva rural del Perú, a pesar de que se sostenga en algunas investigaciones que la brecha de acceso a estos activos se ha reducido en los últimos años.

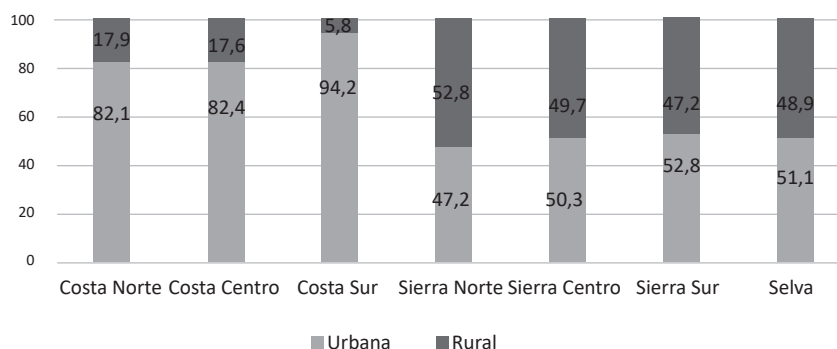
La prueba Chi cuadrado (Tabla 6) permite determinar que la relación entre el área de residencia y el dominio geográfico es significativa a un nivel de confianza del 95 %. Asimismo, con respecto al grado de relación, los coeficientes de correlación muestran valores significativos de alta relación, es decir, pueden mostrar cómo va determinando en alguna medida su influencia en la condición de pobreza que posee el jefe de hogar y cómo esta entrará a tallar dentro la probabilidad que el individuo obtenga dicha condición.

Tabla 6  
Valores y significancia de la prueba Chi cuadrado y medidas simétricas

Prueba Ch cuadrado		Medidas simétricas		
		Coef. de correlación	Valor	Ap prox. Sig.
Pearson Chi cuadrado	447401	Phi	0,586	0,000
		V de Cramer	0,594	0,000
Asymp. Sig. (2-sided)	0,000	Coeficiente de contingencia	0,542	0,000

Fuente: Elaboración propia.

Gráfico 5  
Perfil del jefe de hogar según área y dominio geográfico en estado de no pobreza



Uno de los indicadores que muchos consideran de suma importancia al analizar la pobreza es el de los años de estudios del jefe de hogar. Dada la heterogeneidad y dispersión de las poblaciones en estudio, el tomar como valor del promedio de años



de estudios los valores obtenidos, resulta erróneo. En tal sentido, se ha optado por la mediana como mejor medida de ajuste a esta variable.

Para el año 2016 podemos observar (Tabla 7) que el activo de capital humano referente a los años de estudios del jefe de hogar es bajo, basta solo con observar que para el estado de pobreza la mediana es de 6 años de estudio, pequeña en comparación con la mediana de los jefes de hogar en condición de no pobreza que es de 12 años de estudios.

Tabla 7  
**Cuadro de años de estudio del jefe de hogar según condición de pobreza**

Años de estudio	Estado de pobreza	
	Pobre	No pobre
Media	5,12	11,18
Promedio sin 5% VE	5,98	12,04
Mediana	6	12
Desviación estándar	4,68	5,24
Asimetría	0,586	- 0,347
Kurtosis	- 0.769	- 0,876

Fuente: ENAHO 2016-IV. Elaboración propia.

Al aplicar una prueba de diferencia de medias, se encontró un valor pivotal de  $t = -11,245$  y una significancia de 0,000, lo que no permite concluir a nivel poblacional que existe una diferencia de medias de años de estudio según su condición de jefe de hogar pobre y no pobre. No debemos además olvidar que estamos considerando su conducta desde la idea de un corte transversal hecho en el tiempo y deben ser consideradas sus implicancias respectivas, en caso se decidiese probarla con alguna otra variable y analizar su comportamiento.

## Modelo logit

Como se ha mencionado en la sección correspondiente a la metodología, con base en regresiones logit, se evalúa en primer lugar la probabilidad de que un individuo caiga en condiciones de pobreza como función de los niveles educacionales del jefe de hogar.

En la tabla 8 se presenta la estimación de la primera ecuación econométrica por medio de una regresión logit para todo el país, y se observa el impacto de la educación sobre la probabilidad de ser pobre.

Tabla 8  
**Modelo logit 1, Pobreza y educación de los jefes de hogar, 2016**

Nivel educativo	B	S.E.	Wald	df	Sig.	Exp(B)	Diferencial
Sin nivel (referencial)						43.2(*)	
Primaria incompleta	- 0,312	0,030	14,537	1	0,001	0,097	
Primaria completa	- 0,498	0,034	18,488	1	0,000	0,185	18,5
Secundaria incompleta	- 0,678	0,038	10,245	1	0,000	0,239	
Secundaria completa	- 0,845	0,038	12,987	1	0,001	0,274	8,9
Sup. No Univ. Incompleta	- 1,107	0,065	15,549	1	0,000	0,318	
Sup. No Univ. Completa	- 1,321	0,052	18,156	1	0,002	0,364	9
Sup. Univ. Incompleta	-1,637	0,076	21,834	1	0,000	0,381	
Sup. Univ. Completa	-1,749	0,068	18,954	1	0,000	0,415	14,1

(\*) Probabilidad estimada. Fuente: Enaho 2016-IV. Elaboración propia.

En la tabla se muestra que en el año 2016 si el jefe de hogar no posee ningún nivel educativo, la probabilidad de que el hogar sea pobre es de 43,2 % si se mantiene el resto de las variables constantes. No obstante, a medida que el individuo adquiere mayor educación, esta probabilidad de reduce en 18,5 puntos porcentuales, en el caso de primaria completa; 27,4 puntos porcentuales en el caso de secundaria completa; tener una carrera técnica reduce en 36,4 puntos porcentuales, hasta 41,5 puntos porcentuales, en el caso de un individuo con educación universitaria completa. También se observa que el impacto marginal de culminar un nivel educativo es estrictamente creciente en valor absoluto.

Tal es así que culminar la primaria reduce la probabilidad de ser pobre en 18,5 puntos, culminar la secundaria (con respecto a solo tener primaria completa) reduce la probabilidad en 8,9 puntos. Sin embargo, para analizar el impacto de la educación superior, se debe diferenciar por tipo: no universitaria o universitaria. En el primer caso, culminar la educación superior no universitaria implica una reducción de la probabilidad de ser pobre de 9 puntos, mientras que culminar la educación superior universitaria significa una reducción de la probabilidad en 14,1 puntos. Ambos efectos marginales se calcularon tomando como referencia la probabilidad de ser pobre para un individuo con educación secundaria completa. Los resultados muestran que culminar la educación primaria tiene el mayor impacto en términos de reducir la probabilidad de ser pobre, seguido de la educación superior universitaria, educación superior técnica y, en último lugar, la educación secundaria.

Con relación con el segundo modelo logit, queda demostrada la importancia de la educación en la probabilidad de ser pobre, pero de que por sí sola no podrá exhibir retornos positivos si es que no se consideran algunos aspectos como son los aspectos demográficos, laborales, patrimonio, geográficas y de vivienda.

La variable dependiente es el nivel de pobreza, que puede tomar dos valores: pobre y no pobre. Las variables independientes a considerar se presentan en el Anexo 1, en el cual se presentan las dimensiones y variables independientes en el modelo explicativo.

De la tabla 9, se observa que:

- Si el jefe de hogar no posee ningún nivel educativo, la probabilidad de que el hogar sea pobre aumenta en 2,168 veces si se mantiene el resto de las variables constantes. A medida que el individuo adquiere mayor educación, esta probabilidad de ser pobre se reduce de manera similar a como se observó en el primer modelo logit.
- Si el jefe de hogar vive en la sierra urbana o en la selva urbana, la probabilidad de que el hogar sea pobre aumenta en 1,614 y 1,352 veces, respectivamente.
- Si el jefe de hogar ha estudiado en un colegio estatal, la probabilidad de que el hogar sea pobre aumenta en 1,475 veces.
- Si el jefe de hogar tiene la categoría ocupacional de obrero, la probabilidad de que el hogar sea pobre aumenta en 1,604 veces.
- Si el jefe de hogar trabaja en una empresa que tiene menos de 100 trabajadores, la probabilidad de que el hogar sea pobre aumenta en 1,582 veces.

En cuanto a la variable número de miembros del hogar, se tiene que por cada unidad que aumente el número de personas por hogar hace que la probabilidad de que el hogar sea pobre aumente en 1,815 veces.

- Si el jefe de hogar tiene el estado civil de conviviente, la probabilidad de que el hogar sea pobre aumenta en 1,442 veces.
- Si la vivienda del jefe de hogar no tiene servicios higiénicos, la probabilidad de que el hogar sea pobre aumente en 1,262 veces si se mantiene el resto de las variables constantes.

Por otro lado, las variables, edad, vivir en la costa urbana o selva rural, tener experiencia laboral, tener años de estudio, tener ocupación secundaria, la cantidad de perceptores en el hogar, la vivienda cuente con servicios higiénicos y que tenga agua son factores que disminuyen la probabilidad de que el hogar sea pobre. Así, si el jefe de hogar vive en la costa urbana o en la selva rural, disminuye la probabilidad de ser pobre en 0,35 y 0,34 veces, respectivamente, si se mantienen constantes el resto de las variables.

Tabla 9  
**Modelo logit 2. Variables de la ecuación relacionada  
 con la pobreza del jefe de hogar, 2016**

Variables	B	S.E.	Wald	df	Sig.	Exp(B)
SINNIVEL	0,863	0,206	13,621	1	0,000	2,168
PRIMARIA	- 0,212	0,035	13,896	1	0,000	0,043
SECUNDARIA	-0,273	0,042	14,664	1	0,000	0,098
SUPNOUNIV	- 0,328	0,049	11,278	1	0,000	0,114
SUPUNIV	- 0,644	0,068	12,755	1	0,001	0,214
EDAD	-0,437	0,059	35,628	1	0,000	0,978
SEXFEM	-0,265	0,074	34,762	1	0,000	0,989
COSURB	-1,034	0,123	42,855	1	0,000	0,351
SIEURB	0,670	0,106	35,972	1	0,000	1,614
SELURB	0,612	0,136	48,476	1	0,000	1,352
SELRUR	-0,736	0,145	44,842	1	0,000	0,304
EXPLAB	-0,245	0,095	51,322	1	0,000	0,134
AÑOSEST	-0,227	0,051	24,652	1	0,000	0,985
COLEST	0,368	0,123	14,655	1	0,000	1,475
CUOBRERO	0,396	0,106	18,544	1	0,000	1,604
TEMENOS100	0,462	0,112	9,586	1	0,000	1,582
OCUSEC	- 0,302	0,097	10,386	1	0,004	0,081
MIEMHOG	0,613	0,083	192,034	1	0,000	1,815
PERCEPHOG	-0,425	0,091	72,962	1	0,000	0,954
CONVIVIENTE	0,348	0,103	14,744	1	0,000	1,442
SEHIDVIV	-0,405	0,123	21,459	1	0,000	0,832
NOSERVHIG	0,327	0,095	12,893	1	0,001	1,262
AGUAVIV	-0,204	0,094	7,552	1	0,005	0,763
CONST	0,963	0,563	6,285	1	0,037	2,538

Fuente: Enaho 2016-IV. Elaboración propia.

## Discusión de resultados

El análisis de tablas cruzadas nos ha permitido demostrar que la pobreza se encuentra asociada a niveles bajos de educación alcanzados por el jefe de hogar (relación altamente significativa), tal es así que los jefes de hogar sin nivel educativo (64,6 %) y nivel educativo primario (45,8 %) se encuentran en condición de pobreza, respectivamente, en contraste con una tendencia creciente mucho mayor y estado de no pobreza en la que se encuentran aquellos jefes de hogar que tiene niveles como el universitario y el posgrado, con 93,5 % y 99,97 %, respectivamente. Como indican la mayoría de autores en relación al tema, que a más educación la probabilidad de ser pobre debe ser menor, ya que a mayor educación del jefe de hogar, mayor nivel de ingreso (y por lo tanto, menor probabilidad de ser pobre). El jefe de hogar no solamente sería más productivo (más educado) sino que ganaría más.

También se ha encontrado que existe relación significativa entre el estado civil y el sexo del jefe de hogar en condición de pobreza. Se observa un porcentaje alto de jefes de hogar hombres que tienen pareja (casados o convivientes), en contraste con jefes de hogar mujeres sin pareja (viuda, divorciada, separada o soltera). En general, esto podría reflejar que los hogares liderados por mujeres acusan una situación de mayor vulnerabilidad, pues el grueso de los jefe de hogar sin pareja son mujeres.

Asimismo, se ha encontrado que existe una relación altamente significativa a un nivel de confianza del 95 % entre el área de residencia y el dominio geográfico. Se observa que la mayoría de jefes hogar pobre se encuentran en la costa norte, costa centro y costa sur en el área urbana, y que no existe una diferencia significativa entre los jefes de hogar en la sierra y selva ya sea en el área urbana o rural. El hecho de vivir en zonas rurales reduce la probabilidad de ser pobre en comparación con la costa. Aunque los ingresos por trabajo en la costa son superiores a los de las otras regiones, el monto de la canasta básica de consumo en las zonas rurales es casi 50 % menor que en las urbanas. Ello podría explicar el por qué vivir en la sierra o selva rural reduce la probabilidad de ser pobre.

En relación con los años de estudio, se encontró una mediana de 6 años para los jefes de hogar pobre y de 12 años para los jefes de hogar no pobre. En el aspecto estadístico se concluye a nivel poblacional que existe una diferencia de medias de años de estudio según la condición de pobre y no pobre del jefe de hogar.

Con respecto a la aplicación del modelo logit 1, en el que se analiza la relación del estado de pobreza del jefe de hogar con los niveles de estudio alcanzados, se encontró en general, que mientras mayor es el nivel que aprobó en sus estudios, mayores son las reducciones en la probabilidad de ser pobre, el efecto marginal de tener una profesión técnica reduce la probabilidad de ser pobre en 36,4 % y si tiene una profesión universitaria reduce la probabilidad de ser pobre en 41,5 %. Estos resultados concuerdan con los hallados por Salazar (2012) para quien la tendencia es que a mayores niveles de escolaridad, mayor es la reducción en la probabilidad de ser pobre. Tal como lo indica Salazar, el hecho de tener mayor educación no solo reduce la probabilidad de ser pobre, sino que, como lo sostiene Glewwe y Hall (1992),

un mayor nivel educativo reduce la vulnerabilidad, ya que permite que las personas más calificadas sean favorecidas por los cambios tecnológicos y se adapten mejor.

En relación con el modelo logit 2, encontró que si el jefe de hogar no posee ningún nivel educativo, si vive en la sierra urbana o en la selva urbana, si ha estudiado en un colegio estatal, si tiene la categoría ocupacional de obrero, si trabaja en una empresa que tiene menos de 100 trabajadores, si el número de miembros del hogar es alto, si tiene el estado civil de conviviente, y si la vivienda del jefe de hogar no tiene servicios higiénicos, aumenta la probabilidad de que el hogar sea pobre. Mientras que las variables, edad, vivir en la costa urbana o selva rural, tener experiencia laboral, tener años de estudio, tener ocupación secundaria, la cantidad de perceptores en el hogar, la vivienda cuente con servicios higiénicos y tenga agua son factores que disminuyen la probabilidad de que el hogar sea pobre.

El hecho de vivir en zonas rurales reduce la probabilidad de ser pobre en comparación con la costa. Esto debido a que el monto de la canasta básica de consumo en las zonas rurales es casi 50 % menor que en las urbanas. Ello explicaría por qué vivir en la selva rural reduce la probabilidad en 3,04 veces.

Se demuestra que el incremento de la probabilidad de ser pobre se asocia a niveles salariales reales que están asociados a la categoría ocupacional de ser un empleado y un obrero, así como el tamaño de la empresa. El resultado que mide la variable estado civil que fue incluida con el propósito de controlar el hecho que dos personas juntas son menos vulnerables que una sola, donde los dos puedan trabajar con lo cual los ingresos serían mayores, es lo contrario a lo esperado, vale decir que aumenta la probabilidad de ser pobre tal como lo demuestra el hecho de que el jefe de hogar bajo la condición de convivencia hace que sea más vulnerable que una sola. Este es un resultado similar al encontrado por Shack (1999). También se confirma que los pobres tienden a vivir en hogares relativamente más numerosos. El tamaño de la familia tiene un efecto positivo sobre la probabilidad de pertenecer a un hogar pobre. Asimismo, el no contar con servicios higiénicos está asociados a hogares más pobres, resultado que es consistente con la mayoría de estudios realizados en esa dirección (Quispe, 2000; Yamada, 2007).

Con respecto al hecho de que el jefe de hogar sea de sexo femenino también reduce la probabilidad de ser pobre o explicado de otra manera el hecho de que el jefe de hogar sea hombre muestra un efecto positivo sobre la probabilidad de ser pobre.

También se infiere que un jefe de hogar que tenga experiencia laboral, así como un adecuado número de años de estudio hace que su hogar sea menos vulnerable a la pobreza, ya que estas variables también están asociadas con altos niveles de ingreso.

Si el jefe de hogar ha estudiado en un colegio estatal, la probabilidad de que el hogar sea pobre aumenta en 1,475 veces. Este dato confirma lo encontrado por Moron (2009) donde los individuos que concluyeron la educación básica en centros privados son menos propensos a ser pobres, a diferencia de aquellos que estudiaron en centros estatales.

## Conclusiones

El análisis de tablas cruzadas nos ha permitido demostrar que existe una relación altamente significativa entre la pobreza y los niveles de educación alcanzados por el jefe de hogar, tal es así que los jefes de hogar sin nivel educativo (64,6 %) y nivel educativo primario (45,8 %) se encuentran en condición de pobreza, respectivamente; en contraste con una tendencia creciente mucho mayor y estado de no pobreza en la que se encuentran aquellos jefes de hogar que tiene niveles como el universitario y el posgrado, con 93,5 % y 99.97 %, respectivamente.

Existe una relación muy significativa entre el estado civil y el sexo del jefe de hogar en condición de pobreza, ya que es mayor el porcentaje de jefes de hogar hombres que tienen pareja (casados o convivientes), en contraste con jefes de hogar mujeres sin pareja (viudas, divorciadas, separadas o solteras), lo cual mostraría que los hogares liderados por mujeres acusan una situación de mayor vulnerabilidad, pues el grueso de los jefes de hogar sin pareja son mujeres.

Existe una relación altamente significativa a un nivel de confianza del 95 % entre el área de residencia y el dominio geográfico, ya que la mayoría de jefes de hogar pobre se encuentran en el área urbana de la costa norte, centro y sur del país, y no existe una diferencia significativa entre los jefes de hogar en la sierra y selva, tanto en el área urbana como en el área rural.

Existe una diferencia estadísticamente significativa entre el promedio de años de escolaridad según la condición de pobre y no pobre del jefe de hogar.

A través del primer modelo logit se ha encontrado que la educación –por sus efectos sobre la productividad y la generación de ingresos– es un instrumento clave en la política de superación de la pobreza, pues a mayor nivel de estudios aprobado, mayor es la reducción en la probabilidad de ser pobre. También es importante resaltar la diferencia de impacto para distintos grados educativos. En especial se observa que concluir la universidad significa reducir la probabilidad de ser pobre en 14,1 puntos porcentuales, con respecto a un individuo que concluye secundaria, y concluir una carrera técnica significa reducir la probabilidad de ser pobre en 9 puntos porcentuales con respecto a un individuo que concluye secundaria.

En el segundo modelo logit queda demostrada la importancia de la educación en la probabilidad de ser pobre, pero de que por sí sola no podrá exhibir retornos positivos si el diseño de las políticas públicas no son eficientes y no se consideran algunos aspectos como son los aspectos demográficos, laborales, patrimoniales, geográficos y de vivienda, que como hemos analizado y destacado en varios cuadros, son los que también explican la probabilidad de ser pobre. Así tenemos que la probabilidad de que el hogar sea pobre aumenta si el jefe de hogar no posee ningún nivel educativo, si vive en la sierra urbana o en la selva urbana, si ha estudiado en un colegio estatal, si tiene la categoría ocupacional de obrero, si trabaja en una empresa que tiene menos de 100 trabajadores, si el número de miembros del hogar es alto, si tiene el estado civil de conviviente y si la vivienda del jefe de hogar no tiene servicios higiénicos.

## Referencias bibliográficas

- Bernal, E. (2015). Impacto de la educación sobre la pobreza y la desigualdad en Bolivia. *Revista Estudios de Políticas Públicas*.
- Chacaltana, J. (2006). ¿Se puede prevenir la pobreza? Lima: Consorcio de Investigaciones Económicas y Sociales.
- Instituto Nacional de Estadística e Informática (2017). *Evolución de la Pobreza Monetaria 2007-2016*. (Informe Técnico). Lima: INEI.
- Instituto Nacional de Estadística e Informática (2000). Metodología para la medición de la pobreza en el Perú. *Metodologías Estadísticas*, Año 1-N°02
- Morón, E., Castro, J., y Sanborn, C., (2009). Ayuda a las reformas a entregar Crecimiento Inclusivo en el Perú. En L. Rojas (Ed.), *América Latina: un problema económico. Marco de crecimiento aplicado a Brasil, Colombia, Costa Rica, México y Perú*, (pp. 236-293). Lima: Centro para el Desarrollo Global.
- Ramos, V. (2001). *La Educación y la circularidad de la pobreza en el Perú*. Lima. INEI.
- Salazar, V., Quispe, V. y Choque, R. (2012). Educación y Movilidad Social en el Perú. En: H. Contreras (Ed.). *Evidencia para políticas públicas en educación superior*. Lima: Educativo. Ministerio de Educación.
- Sánchez, C. (2015). *Evaluación del impacto de la Educación sobre la pobreza en Costa Rica*. Costa Rica: Universidad de Costa Rica.
- Shack, N. (1999). *La Educación y la probabilidad de ser pobre en el Perú de hoy. Aplicaciones de un modelo probit de maximoverosimilitud*. En: Instituto Cuanto (Ed). *Pobreza y Economía Social. Análisis de una encuesta. ENENVID 1997*. Lima: Instituto Cuanto.
- Verdera F. (2007). *La pobreza en el Perú. Un análisis de sus causas y políticas para enfrentarla*. Lima: Instituto de Estudios Peruanos.
- Yamada, G. y Castro, J. (2007). *Pobreza, Desigualdad y Políticas Sociales en Perú*. Lima: Universidad del Pacífico.
- Yamada, G., y Castro, J., (2008) Gasto Público y Desarrollo Social en Guatemala: Diagnóstico y Propuesta de Medida (Documento de Discusión). Lima:, Centro de Investigación de la Universidad del Pacífico.
- Yamada, G., y Castro, J., (2012). La Educación que Queremos. En B. Seminario, C. Sanborn y N. Alva (Eds.) *Cuando Despertemos en el 2062: Visiones del Perú en 50 años*. (pp. 403–42...). Lima: Universidad del Pacífico.



# Análisis de la pobreza y distribución del ingreso 2009 -2016

Renán Jesús Quispe Llanos

## Resumen

Este trabajo de investigación pretende mostrar dos temas que están relacionados y determinan la condición social de una persona u hogar, además de que inciden en la calidad de vida que tendrá: el ingreso y la pobreza. En efecto, la pobreza es una condición de personas u hogares cuya principal caracterización es de insatisfacción de alguna necesidad. En cambio, el ingreso relacionado con la persona o el hogar, es un recurso que les corresponde a todos, siempre y cuando tenga una fuente que la genera, y cuya magnitud le permite el acceso a la satisfacción de las necesidades humanas o el no acceso para satisfacerlas.

Se presenta, asimismo, un análisis del ingreso, los indicadores de concentración y los factores determinantes del mismo en el periodo 2009-2016. A partir del cual se aplica un modelo semilogarítmico para cuantificar la dependencia del ingreso respecto a las variables mencionadas en un modelo. El estudio se orienta hacia el análisis de la pobreza en el periodo 2009-2016, desde su construcción, a partir de las líneas de pobreza y pobreza extrema, así como las incidencias de pobreza y pobreza extrema. Se estudia también las necesidades básicas insatisfechas y se realiza el análisis integrado de la pobreza. Finalmente, se expone analíticamente los componentes o factores que inciden en la pobreza, como preámbulo de su modelización.

**Palabras clave:** Ingreso, indicadores de distribución del ingreso, factores determinantes del ingreso del hogar; pobreza y pobreza extrema, y factores explicativos de la pobreza.

## Introducción

Uno de los temas sociales de los programas de las organizaciones políticas que aspiran a ocupar el Gobierno, es mejorar las condiciones de vida de la población, las cuales tienen varias formas de caracterización. Un hogar en condiciones inferiores de satisfacer sus necesidades demuestra una menor calidad de vida. Muchos estudios se han preocupado por sistematizar, desagregar y organizar conceptualmente este constructo, así como en sintetizar las diferentes manifestaciones de la pobreza en un indicador que represente esta problemática social. No obstante, podemos considerar inicialmente dos enfoques para analizar la pobreza: el subjetivo (o de autopercepción) y el objetivo. Desde el punto de vista objetivo la pobreza se mide a través del ingreso o gasto de cada uno de los hogares, el cual se compara con dos líneas que delimitan los pobres extremos; los pobres no extremos, y los que no son pobres.

En forma complementaria, por su capacidad de permitir explicar en forma integral las características del nivel socioeconómico de la población, el ingreso constituye un indicador sintético para posicionar a las familias en el nivel que les corresponde; a partir de ello se le puede utilizar como una categoría de estudio para analizar los diferentes perfiles de la población.

En el Perú, la pobreza ha caído en los últimos años. Entre los años 2004 y 2015 se redujo 36,9 puntos porcentuales, al pasar de 58,7 % a 22,7 %, respectivamente. Por área de residencia, la pobreza es muy diferenciada, en el año 2015 el área urbana representa 3,1 veces que la rural. Así, mientras en el área urbana la pobreza alcanzó 14,5 %; en el área rural esta fue 45,2 %. El ingreso Real promedio del decil más bajo creció 36,5 % entre 2009 y 2015, ante un incremento de 6,2 % del decil más alto.

Los trabajos de baja calificación son una de las principales características de las personas pobres, los cuales están estrechamente ligados a su grado de instrucción. Los pobres tienen una baja instrucción o una menor cantidad de años de estudio en comparación con las personas no pobres. La mayoría de los pobres apenas han estudiado la primaria y la tasa que registran aquellos que no han realizado estudios también alta.

Conjuntamente, al realizar el análisis y evolución de la pobreza, se pretende identificar los elementos centrales que explican que un individuo o un hogar sean pobres. En lo concerniente a la distribución del ingreso, es importante examinar algunas propuestas metodológicas con la construcción de ciertos indicadores, que complementen los ya existentes. Estas se presentan bajo dos tipos de perspectiva, la académica y la política, con cifras documentadas que permitan la mejora en el diseño de programas de lucha contra la pobreza.

## Algunos antecedentes

El tema de la medición estándar de la pobreza se aborda en América Latina en la década de los noventa, en las reuniones de trabajo de países liderados por la CEPAL

y el auspicio inicial del BID y del Banco Mundial. Sin embargo, en el Perú, ya en 1975, Richard Webb y Adolfo Figueroa habían realizado un trabajo pionero sobre distribución del ingreso.

López (2010) evalúa los efectos del gasto público y las políticas comerciales sobre la pobreza y la distribución dentro de un marco simultáneo, que considera las interdependencias entre las consecuencias de la liberalización comercial y las políticas de gasto público para la pobreza y la distribución del ingreso.

El Banco Mundial (2009) analiza los cambios en la pobreza a nivel nacional y relaciona crecimiento, desigualdad y reducción de la pobreza. De otro lado, aborda la evolución de la pobreza en el área urbana desde una perspectiva multidimensional y la del área rural mediante el análisis de los patrones de diversificación y especialización de las fuentes de ingreso que permiten a los hogares pobres mejorar sus condiciones de vida. Concluye en su estudio que pese al crecimiento sostenido de los últimos años en el Perú, las regiones presentan diferentes elasticidades de la pobreza al crecimiento, y que por ello la mejora en las condiciones de vida ha sido limitada en el área rural (en especial en la sierra).

Carlos Filgueira, Andrés Peri (2004) precisan que es poco lo que se puede esperar del crecimiento económico como vía primaria para remover los riesgos sociodemográficos, en principio porque el crecimiento en la región es poco y el mismo, contribuye escasamente a disminuir la pobreza y la desigualdad.

La Comisión para la medición de la pobreza (2014) establece que para un estudio profundo de la pobreza se requiere un acercamiento cualitativo a la realidad de los hogares en situación de pobreza y de privación social.

Teitelboim (2003) obtiene un modelo que cumple con los supuestos y todos los test estadísticos requeridos para determinar los factores determinantes de la pobreza tanto en la zona rural como urbana.

Agüero (2008) nos muestra un estudio en el cual determina las variables más relevantes que afectan la pobreza.

Pérez y Rodríguez (2015) llegan a la conclusión de que los componentes de la pobreza multidimensional que se han mantenido en niveles considerables son el nivel educativo del jefe de familia y el déficit calórico por hogar, lo cual puede ser la causa de que la disminución de la incidencia de este enfoque sea menor que en el monetario.

Mendoza, Leyva y Flor (2011) sostienen que la explicación de la desigualdad parece estar en el estilo de crecimiento económico de una economía abierta basada en la exportación de productos primarios y en la incapacidad del Estado para modificar, a través de la política fiscal, la distribución del ingreso generada por el mercado.

Para la OCDE (2016) una de las maneras de aumentar la diversificación económica y la productividad es mejorar la conectividad en el transporte, ello promueve la competitividad y la inclusión, además de que contribuye a reducir las tasas de informalidad laboral. Una buena conectividad puede hacer del Perú un país más

eficiente y garantizar un desarrollo sostenible y más inclusivo. La brecha del Perú con respecto a Alemania (país miembro de la OCDE que muestra los mejores resultados) es un 35% mayor que la del país líder de la región, Chile, y más de 2.5 veces la brecha promedio de los países miembros de la OCDE, dado que el Perú se ha concentrado tradicionalmente en el transporte por carretera y el uso de otros medios de transporte es escaso.

Robles (1997) identifica las características y factores que más influyen en el nivel de ingreso de los hogares, que es a su vez, una de las medidas de la pobreza. El estudio abarca la base de datos correspondiente a la ENAHO98-IV e información complementaria que fue proporcionada por el INEI. Se presenta un modelo de regresión múltiple del ingreso con base en diferentes variables que podrían explicar la distribución de ingreso.

Sierra (1999) identifica los factores que determinan el ingreso de los jóvenes que trabajan en el país y diseña un modelo de regresión que permite explicar dichos ingresos. El investigador se enfoca, principalmente, en los niveles de ingreso de los jóvenes que trabajan por dominio geográfico, área, sexo, edad, nivel de educación y tipo de trabajo.

En síntesis, la pobreza tiene muchas manifestaciones, entre las cuales se puede señalar el bajo nivel de ingreso y el gasto del hogar, alguna necesidad básica insatisfecha, baja posesión de activos del hogar, altas tasas de desnutrición crónica, mayor tasa de mortalidad infantil en el hogar, entre otras; pero la más importante está asociada con el ingreso permanente del hogar. Entre sus principales factores explicativos se puede mencionar el nivel de educación del jefe del hogar, el nivel educativo de sus miembros, el número de perceptores del hogar; el área o lugar de residencia; la edad del jefe del hogar, la tenencia de activos o tierras del jefe del hogar, entre otros. En términos macroeconómicos el desarrollo económico sintetizado en un crecimiento de la economía y una mejora en el ingreso promedio de la población favorece su reducción.

## **Las variables**

### **El ingreso y su medición**

Para fines del presente estudio, el ingreso de los hogares es definido como la suma de ingresos que perciben durante un periodo de tiempo los miembros del hogar, y que puede ser gastado en bienes y servicios para satisfacer las necesidades de consumo de los hogares.

El ingreso familiar es un indicador que sintetiza el acceso a un nivel de vida en el hogar. En el corto plazo genera las condiciones para adquirir bienes y servicios de la canasta familiar en número y calidad, proporcional al nivel de ingreso. En el mediano y largo plazo, a partir de un buen ingreso permanente, facilita la adquisición de activos en el hogar, dotándolo de un mayor confort, que igualmente incide en mayores gastos

por mantenimiento y seguridad. Es decir, es la fuente de financiamiento o de acceso de los bienes y servicios que constituyen la canasta familiar. Un mayor nivel de ingreso genera un mayor poder adquisitivo de la población facilitándole en primer lugar una mejora en el nivel de Bienestar, como consumidor. En los estratos de mayor ingreso permite una capacidad de ahorro o de inversión en activos para el hogar en primera instancia y la posibilidad de participar en la actividad productiva como propietario de los medios de producción.

## **Acerca de la pobreza**

De una manera general se dice que la pobreza existe en una sociedad cuando una o más personas muestran un nivel de bienestar inferior al mínimo necesario para la sobrevivencia. Esta definición tiene muchas connotaciones y en términos socioeconómicos, la situación de pobreza suele determinarse a partir de la consideración de una serie de factores –principalmente el nivel de ingresos– que definen la posibilidad de cubrir necesidades elementales como la alimentación, la vivienda; el acceso a bienes y servicios, a la educación y a la salud. Una persona con todos estos aspectos satisfechos se ubicaría por encima de la línea de bienestar, es decir, fuera de una situación de pobreza.

Otra manifestación de la pobreza es a través de la insatisfacción de al menos una de las necesidades básicas de un hogar. En particular, para efectos de evaluación y formulación de políticas sociales, es necesario saber, además, cuántos pobres hay en la sociedad, qué características tienen y cuáles son sus causas.

## **Manifestaciones de la pobreza**

Salama y Destremau (2012) sostienen que la pobreza afecta más concretamente a los niños. Cuando en una ciudad aumenta la pobreza se incrementa el riesgo de mayor desnutrición crónica de los niños menores de 5 años, así como de la mortalidad infantil. Algunos niños dejan de estudiar por dedicarse a trabajar para aportar en el ingreso familiar, y en esas condiciones se puede incrementar la delincuencia juvenil. El factor educativo es determinante y la composición de familias numerosas conlleva a elaborar estrategias para completar la insuficiencia económica.

Precisan que “el individuo que nace pobre, tiene probabilidades de seguir siéndolo toda su vida y así dará a luz a niños pobres”. Por tanto:

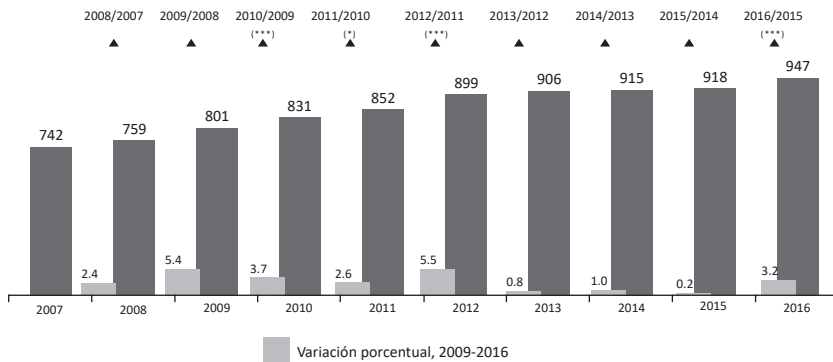
Se genera así la circularidad de la pobreza. El niño pobre sufre dificultades, siendo más álgidos en sociedades desiguales. Entonces el niño debe trabajar para ayudar a la familia a sobrevivir y consecuentemente la asistencia al centro educativo se complica generando abandono escolar.

## Análisis comparativo del ingreso, gasto promedio real per cápita mensual, 2009-2016

En cuanto a la evolución del ingreso real promedio mensual per cápita, se aprecia un alza sostenida del ingreso, que pasó de 742 soles en el año 2007 hasta alcanzar un promedio per cápita real de 947 soles, que representó un incremento de 3,2 %, respecto al promedio alcanzado el año 2015, y una variación promedio anual entre los años 2007 y 2016 de 2,7 %.

Gráfico 1

### Perú: evolución del ingreso real promedio per cápita mensual, 2007-2016 (Soles constantes base=2016 a precios de Lima Metropolitana)



\*Diferencia significativa ( $p < 0.10$ ) \*\*Diferencia altamente significativa ( $p < 0.05$ )

\*\*\*Diferencia muy altamente significativa ( $p < 0.01$ )

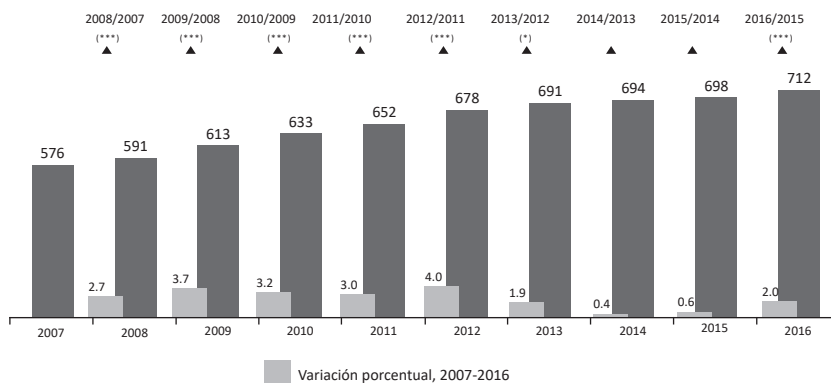
Fuente: INEI - Encuesta Nacional de Hogares, 2007-2016

En lo referente al gasto real promedio per cápita mensual, igualmente se registra un alza sostenida sustentada en el aumento del poder adquisitivo de la población, como se observó líneas arriba. En el año 2007, el nivel del gasto fue 576 soles y llegó a los 712 soles en el año 2016, es decir, experimentó una variación promedio anual de 2.4%. Por otro lado, en el año 2016 se registró un incremento de 2,0% (equivalente a S/. 14 soles) respecto al promedio del gasto real per cápita alcanzado en el año 2015 (gráfico 2).

En consecuencia, esa mejora en el poder adquisitivo de los ingresos, no solo permitió una mejora en el gasto, sino que además se pudo incrementar el ahorro real. Las estadísticas macroeconómicas nos muestran una elevación del consumo privado. Asimismo, se observó una reducción de la pobreza, aunque de una manera muy ligera.

Gráfico 2

**Perú: evolución del gasto real promedio per cápita mensual, 2007-2016**  
**(Soles constantes base=2016 a precios de Lima Metropolitana)**



\*Diferencia significativa ( $p < 0.10$ ) \*\*Diferencia altamente significativa ( $p < 0.05$ )  
 \*\*\*Diferencia muy altamente significativa ( $p < 0.01$ )  
 Fuente: INEI - Encuesta Nacional de Hogares, 2007-2016

## Análisis de la distribución del ingreso en el año 2016

### Distribución del ingreso a nivel nacional en el año 2016

En el año 2016, de acuerdo a los resultados de la ENAHO en el Perú, los hogares percibieron mensualmente en promedio la suma de 2 611.23 soles (aproximadamente, 792 dólares) y en términos per cápita, la suma de 826 soles por mes.

El ingreso per cápita del decil de mayores ingresos es 23 veces superior al ingreso per cápita del decil de menores ingresos, y mientras un hogar del primer decil percibe en términos per cápita 129 soles, en promedio, el último decil percibe 2 931 soles.

El ingreso per cápita mensual de los hogares por deciles para el año 2016, muestra que los primeros siete deciles obtienen un ingreso per cápita inferior al promedio nacional (826 soles), el cual se encuentra fuertemente influenciado por el ingreso de los tres últimos, donde el último decil alcanza 2 931 soles; es decir, la distribución del ingreso tiene una asimetría positiva, sesgada hacia la derecha.

Por el lado del promedio anual de los ingresos corrientes per cápita, gasto corriente per cápita y ahorro corriente estimados de los hogares a nivel nacional, cada uno de ellos –según deciles de ingreso– presenta a los ahorros corrientes que alcanzan la cifra promedio de 1 580.25 soles, que representa el 4.97% del ingreso corriente del año 2016. Asimismo, se confirma los cambios en los patrones de consumo al variar de un nivel a otro, así como la alta propensión marginal a consumir en los primeros deciles, en respuesta a mayores niveles de ingreso, lo que repercute en un desahorro que se eleva a mayores niveles de ingreso hasta el cuarto decil (Anexo 2).

Con respecto al nivel de desahorro, en el año 2016 se percibe una mejora respecto al año 2009. En efecto, se observa que en el año 2016 los desahorros se dan hasta el séptimo decil, en cambio en el 2009 era hasta el octavo. Asimismo, los niveles de desahorro en el 2016 igualmente eran significativos hasta el sexto decil; en cambio en el séptimo, el saldo negativo es de 487.80 soles. Se puede advertir las menores propensiones marginales a consumir en los niveles superiores (Anexo 3).

Es posible obtener una medida que resume el grado de desigualdad en la distribución del ingreso a nivel nacional a través del coeficiente de GINI que alcanza un valor de 0,46 para el 2016, ello confirma la reducción registrada en el año 2015 (Anexo 4).

Con respecto a los indicadores de desigualdad, tanto de Gini, índice de concentración de Theil, varianza del logaritmo y coeficiente de variación en los años 2009 y 2016 a nivel nacional nos muestran en conjunto que la desigualdad del ingreso en el mediano plazo se reduce, lo que se corrobora con un aumento de la clase media en el Perú (Anexo 5).

## Metodología

### Los factores determinantes del ingreso en modelo

Para el año 2016 se aplica un modelo transversal, cuyo análisis se va a realizar de la misma forma para todas las variables determinadas de manera histórica. El modelo se aplica de acuerdo con algunos textos de la bibliografía: *Factores determinantes de los ingresos del hogar* y *Determinantes del ingreso de los jóvenes que trabajan*.

La ecuación del modelo semilogarítmico, es:

$$\ln Y = a + bX + m \quad (1)$$

Donde:  $a = a_0 - b$

De (1)  $(1/Y) (dY/dX) = b$

Así pues, la tasa de cambio proporcionada en Y por un cambio unitario en X es constante e igual a b. La función solo está definida para valores positivos de Y. Si se omite la perturbación en (1), podemos reescribir la función como:

$$Y = e^a + bX \quad (2)$$

Donde:

Y: variable dependiente Ingreso

b: coeficientes a estimar

a: valor constante

X: variables independientes

e: base del logaritmo natural, aproximadamente 2,718



**Tabla 1**  
**Distribución de las variables que determina el ingreso de los jefes del hogar, según factores**

Factor	Variables
1. Educativo - personal	Año de estudio
	Sexo
	Edad
	Estado civil
2. Laboral	Categoría ocupacional
	Rama de actividad
	Tamaño centro laboral
	Perceptores del hogar
3. Geográfico y de vivienda	Región
	Combustible para cocinar
	Total de dormitorio
	Servicio higiénico
4. Tenencia de activos físicos	Material de piso
	Auto para uso particular
	Horno microondas
	Computadora
	Refrigeradora
	Televisor de color

*Fuente: Elaboración propia*

Se ha identificado las variables económicas, sociales y demográficas que más influyen en la determinación del ingreso de los jefes del hogar que trabajan.

A continuación, se presentan el modelo de regresión del ingreso de los jefes del hogar que trabajan:

Tabla 2  
Modelo de regresión del ingreso de los jefes del hogar

Dependent Variable LN\_Y\_  
Method Least Squares  
Date 07/02/17 Time: 22.00  
Sample 1 35785  
Included observations: 35288

Variable	Coefficient	Std Error	t-Statistic	Prob.
TV	0.681066	0.013786	49.40120	0.0000
AÑO DE ESTUDIO	0.147361	0.002077	70.93448	0.0000
SEXO	0.373386	0.013162	28.36806	0.0000
EDAD	0.026290	0.000328	80.25182	0.0000
ESTADO CIVIL	0.126574	0.003939	32.13295	0.0000
CATEGORÍA OCUPACIONAL	0.111387	0.005690	19.57619	0.0000
RAMA DE ACTIVIDAD	0.059752	0.002934	20.36463	0.0000
TAMAÑO CENTRO LABORAL	0.064064	0.004459	14.36802	0.0000
PERCEPHO	0.136261	0.005369	25.37756	0.0000
REGIÓN DOMINIO	0.084821	0.002220	38.20737	0.0000
COMBUSTIBLE PARA COCINAR	-0.070512	0.003158	-22.33062	0.0000
TOTAL DE DORMITORIO	0.067974	0.003621	18.77090	0.0000
SERVICIO HIGIÉNICO	0.019598	0.002656	7.378958	0.0000
MATERIAL DE PISO	0.204975	0.004244	48.30084	0.0000
AUTO	0.417820	0.021300	19.61624	0.0000
MICROONDAS	0.202217	0.017340	11.66186	0.0000
COMPUTADORA	0.190745	0.014561	13.09960	0.0000
REFRIGERADORA	0.329474	0.013690	24.06630	0.0000
R-squared	0.241663	Mean dependent var		5.905449
Adjusted R-squared	0.241298	S.D. dependent var		1.115346
SE of regression	0.971505	Akaike info criterion		2780570
Sum squared resid	332883.63	Schwarz criterion		2.784891
Log likelihood	-49042.38	Hannan-Quinn criter.		2781946
Durbin-Watson stat	1.780357			

## **Análisis de resultados**

La investigación confirma que la educación y las características personales tienen un efecto positivo sobre el ingreso de los jefes del hogar que trabajan y es el factor fundamental de la determinación del ingreso, el sexo y los años de estudio que tiene el jefe del hogar. Esto quiere decir que por cada año de estudio del jefe del hogar, el nivel de ingreso per cápita aumentará en 14.74%; y lo hará en 37.34% por el factor sexo. Respecto al factor laboral las principales variables que determinan son la categoría ocupacional y los perceptores del hogar. Estos afectarán al nivel del ingreso por cada categoría y perceptores, de modo que aumenten este mismo en 11.11% y 13.62%, respectivamente.

Según el modelo de regresión para los factores geográfico y de vivienda la variable más importante por considerar en este factor es el material de piso, el cual alcanza un coeficiente estimado de 20.50% respecto de las demás variables que representan el factor mencionado.

También contribuye de una manera importante al ingreso de los jefes del hogar, la tenencia de activos físicos, que son los que tienen mayor coeficiente y relevancia en el modelo. Son los más importantes, la televisión a color y el automóvil para uso particular, los cuales se poseen cuando aumenta el nivel de ingreso per cápita de los jefes del hogar en 68.11% y 41.78% por cada unidad.

Teniendo en cuenta el modelo de regresión se puede plantear que los ingresos de los jefes del hogar están fundamentados por la tenencia de los activos tangibles que tengan, que son productos que se adquieren sustentados por un mayor ingreso; y su factor tanto educativo como personal, en el ámbito de año de estudio y sexo.

En conclusión, se está de acuerdo con la insistencia de las nuevas teorías del crecimiento, en la importancia de la educación y de todos los elementos que son parte del capital humano; pero, por lo menos, para el caso de los jefes del hogar que trabajan en el país, la educación es una de las determinantes del ingreso. Asimismo, la posesión de activos se basa en los mayores ingresos.

## **Métodos para medir pobreza**

Existen diversos métodos para identificar a los pobres, ya sea por ejemplo los métodos que miden la situación de los hogares respecto a sus características, el medio donde se desenvuelven y el acceso a los servicios básicos; así como los métodos que miden la pobreza según el poder adquisitivo de los hogares mediante líneas de pobreza.

Según Feres y Mancero (2001): “la teoría no permite elegir un método por sobre los demás, ya que todos tienen defectos y virtudes; más bien, es la práctica quien ha resuelto la elección de un método, de acuerdo al contexto en el que se aplica”.

A continuación, se muestran algunas definiciones de pobreza así como un breve resumen de las metodologías para su cálculo, ventajas y desventajas, con la finalidad de brindar a los usuarios de esta información, elementos útiles para su interpretación.

## La pobreza monetaria

Definida como la insuficiencia de recursos monetarios para adquirir una canasta de consumo mínima, que sea aceptada socialmente. Por lo que se elige un indicador de bienestar (gasto per cápita) y ciertos parámetros socialmente aceptados (líneas de pobreza total y pobreza extrema para los casos de consumo total y caso de alimentos, respectivamente). Por esto se dice que un hogar es pobre cuando su gasto per cápita es inferior a una Línea de Pobreza (LPt) y un hogar es pobre extremo cuando su gasto per cápita es inferior a una Línea de Pobreza Extrema (LPex).

En referencia a la tasa de pobreza monetaria, es comúnmente el indicador que hace referencia al nivel de vida de la población, esta refleja la capacidad de un hogar para afrontar las exigencias mínimas para vivir; en este sentido el indicador que se utiliza es el gasto per cápita del hogar.

## Método de línea de pobreza

Este método está más centrado en la dimensión económica de la pobreza y utiliza el ingreso o el gasto de consumo como medidas del bienestar. Al determinar los niveles de pobreza se compara el valor per cápita de ingreso o gasto en el hogar con el valor de una canasta mínima denominada línea de pobreza. Este método se complementa con los indicadores de Foster, Greer y Thorbecke (FGT).

## Indicadores Foster, Greer y Thorbecke (FGT)

Se propone una ecuación que contiene el indicador de la “aversión a la desigualdad”, que muestra la importancia que se les asigna a los más pobres en comparación con los que están cerca de la línea de pobreza.

$$P\alpha = \frac{1}{n} \sum_{i=1}^q \left[ \frac{(z - y_i)}{z} \right]^{\alpha}, \text{ donde } \alpha \geq 0$$

Definimos las variables:

$P\alpha$  : Incidencia de la pobreza

$y_i$  : Consumo per cápita (estimado como el total del consumo del hogar dividido para el número de miembros) del individuo  $i$  ( $i = 1, 2, \dots, q$ ), considerando solo a quienes cuyo consumo es inferior a la línea de pobreza

$n$  : Población total.

$q$  : Número total de pobres.

$z$  : Línea de pobreza.

$\alpha$  (alpha): número real no negativo, que puede tomar uno de tres valores: 0, 1, 2.

Cuando  $a = 0$ , esta medida es igual al índice de recuento ( $H$ ): el porcentaje de personas que se encuentran en situación de pobreza monetaria.

Cuando  $a = 1$ , se obtiene la brecha de pobreza, decir, el porcentaje promedio en el cual los pobres deberán aumentar sus ingresos para salir de la línea de pobreza. De los distintos indicadores que forman parte de este grupo, el que más atención suele recibir es  $P2$  (que se obtiene con  $a = 2$ ), que es una medida de la “severidad” de la pobreza, que representa la distribución de los gastos per cápita entre los pobres.

### **Brecha de la pobreza: Poverty Gap Poor (PGP)**

Brecha de producto –llamado también FGT1–, es considerado el indicador que toma en cuenta el déficit de ingreso ( $z - y_i$ ) de los considerados “pobres monetarios”. La brecha de pobreza mide la insuficiencia promedio del consumo de los pobres respecto de la línea de pobreza y toma en cuenta la proporción de la población pobre en la población total. Se obtiene al agregar el déficit de ingreso de la población con respecto a la línea de la pobreza:

$$PG = \frac{\sum_{i=1}^q \left( \frac{z - y_i}{z} \right)}{n}$$

Donde:

$PG$  : Brecha de la pobreza

$z$  : Línea de pobreza

$y_i$  : Ingresos per cápita

$n$  : Población total

### **Indicador de severidad de la pobreza: Foster Greer Thorbecke (FGT2)**

Este indicador refleja el grado de desigualdad al interior de los pobres, mide las carencias en el consumo privado y toma como referencia una determinada línea de pobreza individual, relacionada básicamente con la severidad de la pobreza. Da mayor peso a las brechas de los más pobres:

$$P2 = \frac{\sum_{i=1}^q \left( \frac{z - y_i}{z} \right)^2}{n}$$

Donde:

$P2$  : Severidad de la pobreza

$z$  : Línea de pobreza

$y_i$  : Consumo per cápita (estimado como el total del consumo del hogar dividido para el número de miembros) del individuo  $i$  ( $i = 1, 2, \dots, q$ ), considera solo a quienes tienen un consumo inferior a la línea de pobreza

$n$  : Población total

## La pobreza no monetaria

Los indicadores de pobreza estructural se refieren estrictamente a la caracterización de los hogares en cuanto a su tenencia o calidad de bienes y servicios, o en todo caso, tenencia de capital humano que describa mejor sus condiciones de vida.

### Indicador de necesidades básicas insatisfechas (NBI)

Es una metodología del INEI utilizada para la medición de la pobreza en el Perú. Toma en consideración un conjunto de indicadores relacionados con las características de los hogares en relación con sus necesidades básicas estructurales: vivienda, educación, salud, infraestructura pública, etc. Se define pobre por NBI a aquella población que reside en hogares con al menos una de las siguientes necesidades básicas insatisfechas:

- **Hogares en viviendas con características físicas inadecuadas.** Toma en cuenta el material predominante en las paredes y pisos, así como el tipo de vivienda.
- **Hogares en viviendas con hacinamiento.** Se determina que hay hacinamiento cuando residen más de 3.4 personas por habitación.
- **Hogares en viviendas sin desagüe de ningún tipo.** Porcentaje de vivienda sin desagüe de ningún tipo.
- **Hogares con niños que no asisten a la escuela.** Hogares con presencia de al menos un niño de 6 a 12 años que no asiste a un centro educativo.
- **Hogares con alta dependencia económica.** Porcentaje de la población en hogares con jefe con primaria incompleta (hasta segundo grado) y con: (i) 4 o más personas por ocupado o (ii) sin ningún miembro ocupado.

En el caso del método de las necesidades básicas insatisfechas, el INEI determina el número de ellas en cada hogar y luego presenta la proporción de personas que tienen por lo menos una NBI (pobres) o por lo menos dos NBI (pobres extremos).

### El método integrado

“El método integrado consiste en plantear una nueva metodología de medición de la pobreza, basada en el “cruce” de resultados entre Necesidades Básicas Insatisfechas (NBI) y Costo de las Necesidades Básicas (LP). Las primeras investigaciones en las que se evidencia este “enfoque bidimensional” se encuentran en Beccaria y Minujin (1985) y en CEPAL-DGEC (1988c), con datos de Argentina y Uruguay, respectivamente.

Tabla 3  
**Combinación de resultados del método de líneas de pobreza  
 y de necesidades básicas insatisfechas**

LP ↓	NBI →	<b>Hogares con NBI</b>	<b>Hogares sin NBI</b>
HOGARES POBRES		Hogares en situación de pobreza crónica	Hogares en situación de pobreza coyuntural (reciente) (pauperizados)
HOGARES NO POBRES		Hogares con carencias estructurales (inerciales)	Hogares en condiciones de integración social

Fuente: CEPAL / DGEC (1988c) y otros estudios sobre el tema.

Esta metodología aprovecha la complementariedad entre NBI y LP y se ha constituido en una suerte de estándar para la caracterización de la pobreza en América Latina.

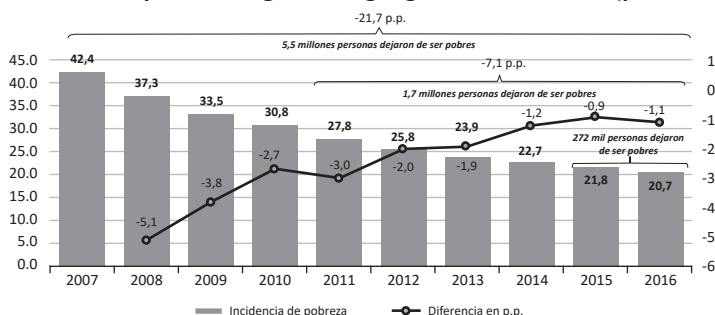
### Análisis histórico de la pobreza por el método del gasto

Se considera que una población vive en condiciones de pobreza monetaria si residen en hogares cuyo gasto per cápita es insuficiente para adquirir una canasta básica de alimentos y no alimentos (vivienda, vestido, educación, salud, transporte, etc.). Una población está en condiciones de pobreza extrema cuando las personas que integran hogares sus gastos per cápita están por debajo del costo de la canasta básica de alimentos.

### Incidencia de la pobreza total

Determina la proporción de la población cuyo consumo se encuentra por debajo del valor de la línea de pobreza. Cabe resaltar que son considerados pobres, aquellos que no pueden cubrir el costo promedio mensual de la canasta básica de consumo, compuesta por alimentos y otros bienes y servicios.

Gráfico 3  
**Incidencia de la pobreza según área geográfica, 2007-2016 (porcentaje)**



Fuente: INEI - Encuesta Nacional de Hogares, 2007 - 2016

Es así que en el periodo comprendido entre los años 2007 y 2016 la tasa de pobreza se redujo en 21.7 puntos porcentuales, lo que representa aproximadamente 5,5 millones de personas que salen de la condición de pobreza, en dicho periodo de tiempo. En los últimos dos años, este comportamiento significó la salida de la pobreza, de aproximadamente 272 mil personas.

### Incidencia de la pobreza extrema

Se considera pobres extremos a la proporción de la población cuyo consumo se encuentra por debajo del valor de la línea de extrema pobreza, cuyo cálculo se mide mediante un índice de conteo de personas que calcula el número de pobres extremos como una proporción de la población total.



Fuente: INEI - Encuesta Nacional de Hogares, 2007 - 2016.

El porcentaje de la población peruana en situación de pobreza extrema pasó de 11.2% (2007) a 3.8% (2016), en cifras absolutas son equivalentes aproximadamente a 1,9 millones de personas que salieron de la pobreza extrema. Del mismo modo, entre los años 2015 y 2016 han logrado salir de la pobreza extrema aproximadamente 80 mil personas, que tenían un gasto inferior al costo promedio de la canasta básica alimentaria.

### Indicadores de la brecha de la pobreza total

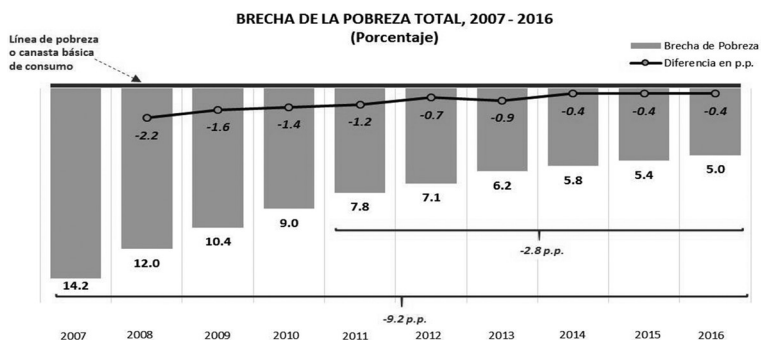
La brecha de la pobreza refleja cuan pobres son los pobres y, por tanto, proporciona una idea de la profundidad de las carencias del consumo que definen una situación de pobreza. Es decir, representa el déficit promedio de consumo de la población para satisfacer las necesidades mínimas de bienes y servicios de todos sus integrantes (que se expresa como proporción de la línea de pobreza), donde el déficit de la población no pobre es cero por definición.



Si se compara la información entre los años 2007 y 2016, a nivel nacional, la brecha de la pobreza disminuyó 9.2 puntos porcentuales, pues pasó de 14.2% en año el 2007 a 5.0% en el 2016.

En el año 2016, la brecha promedio del gasto de los hogares pobres con respecto al costo de la canasta básica de consumo fue de 5.0%. Si se le compara con en el año 2011, que fue de 7,8%, disminuyó en 2.8 puntos porcentuales.

Gráfico 5  
Indicadores de la severidad de la pobreza total

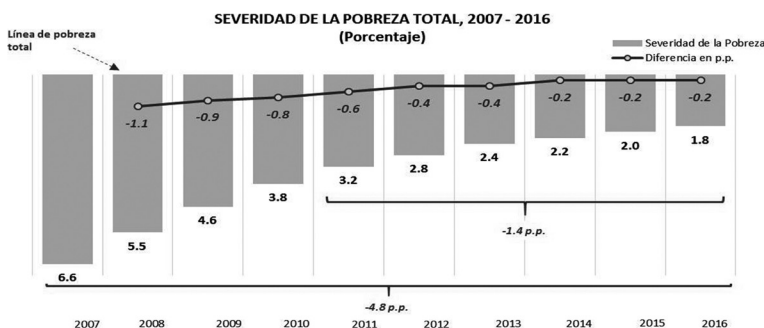


Fuente: INEI - Encuesta Nacional de Hogares, 2007 - 2016.

## Indicadores de la severidad de la pobreza total

Este indicador refleja el grado de severidad o desigualdad de la pobreza al interior de los pobres; medida que depende de la distribución y que pondera —esencialmente— las brechas de pobreza promedio por la población a cada nivel. Además, indica que cuanto más alto sea el valor, existe mayor desigualdad entre los pobres.

Gráfico 6  
Indicadores de la severidad de la pobreza total



Fuente: INEI - Encuesta Nacional de Hogares, 2007 - 2016.

## Análisis de la pobreza por necesidades básicas insatisfechas

El método de las NBI es un método directo para identificar carencias críticas en una población y caracterizar la pobreza, para lo cual generalmente utiliza indicadores que se relacionan directamente con las áreas de necesidades básicas de las personas (vivienda, servicios, educación).

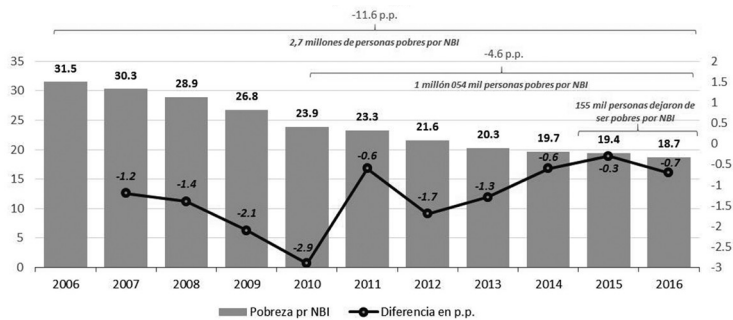
A nivel nacional la tasa de pobreza por NBI en 2016 se ubicó en 18,7 %, lo que evidencia una disminución de 0,7 puntos porcentuales en relación con el año 2015, que fue de 19,4 %, lo que en cifras de población significa 155 mil personas, aproximadamente, dejaron de ser pobres según el método de NBI.

De otro lado, entre los años 2011 y 2016 se observa una disminución de 4,6 puntos porcentuales, que en cifras representa aproximadamente a poco más de un millón de personas que han dejado de ser pobres bajo la aplicación de este método en este periodo.

En el periodo comprendido entre los años 2007 y 2016 la pobreza por NBI se redujo en 11,6 puntos porcentuales (bajó de 31,5 % a 18,7 %), lo que representa aproximadamente 2,7 millones de personas que han salido de la condición de pobreza por el método de NBI.

Gráfico 7

### Evolución de la pobreza por necesidades básicas insatisfechas (Porcentaje)



Fuente: INEI - Encuesta Nacional de Hogares, 2007 - 2016.

## Evolución de la pobreza por el método integrado

Al combinar los resultados producidos por cada método, se genera una tipología que identifica cuatro posibles situaciones, cuya denominación más difundida corresponde a la acuñada por Kztman (CEPAL-DGEC, 1988c):

- Hogares en condiciones de integración social.
- Hogares en situación de pobreza estructural o con carencias inerciales.

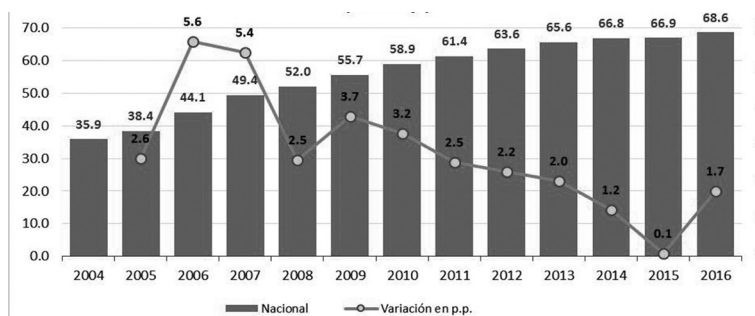
- Hogares en situación de pobreza coyuntural o reciente.
- Hogares en situación de pobreza crónica.

## Población socialmente integrada

Es la población que dispone de un ingreso suficiente para lograr un nivel de vida adecuado y además satisfacen sus necesidades básicas, tienen la capacidad de integrarse plenamente a la sociedad.

Durante el periodo 2004-2016 se observa un incremento representativo de personas con mayor disponibilidad de ingresos que se encuentran por encima de la línea de pobreza; estos ingresos les permiten satisfacer sus necesidades básicas en primer término y ello se traduce en una mejora en su calidad de vida. Este incremento representa en porcentaje el 32,7, es decir, que en el año 2004 fue de 35,39 % y llegó al 68,6 % en 2016.

Gráfico 8  
Población socialmente integrada, 2014-2016  
(Porcentaje)



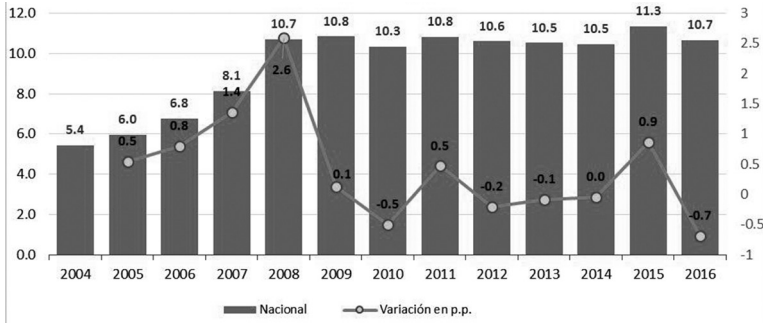
Fuente: INEI - Encuesta Nacional de Hogares, 2007 - 2016.

## Población en pobreza estructural

La población que se encuentra en esta situación es aquella que cuenta con un ingreso suficiente para adquirir los bienes y servicios básicos, pero que no ha logrado mejorar ciertas condiciones de su nivel de vida. Se trata de un conjunto de personas que ha arrastrado ciertas necesidades desde el pasado, las mismas que se han establecido como su estilo de vida.

Del año 2004 al año 2009, la proporción de este tipo de población se incrementó en casi el doble, ya que pasó de 5,4 % a 10,8%. A partir del año 2010, este incremento registra cifras constantes y llega 10,7% de población con pobreza estructural en el año 2016; es decir, alcanza una diferencia entre dichos años de 0,3 puntos porcentuales, pues pasó de 10,3 % (2010) a 10,7% (2016).

Gráfico 9  
**Población socialmente integrada, 2014-2016**  
**(Porcentaje)**

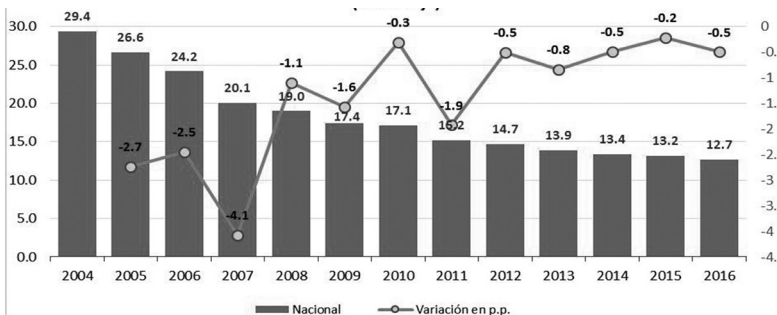


Fuente: INEI - Encuesta Nacional de Hogares, 2007 - 2016.

### Población en pobreza coyuntural

En esta situación de pobreza se encuentran aquellos pobladores que satisfacen sus necesidades básicas pero que tienen un ingreso inferior a la línea de pobreza, porque –de acuerdo con la temporalidad relacionada con cada método de medición– se trataría de hogares en un proceso de “movilidad descendente” (CEPAL, 1988c). Asimismo, es plausible suponer que una reducción reciente de ingresos en estos hogares se traducirá en un menor consumo corriente, y no en un deterioro inmediato de los logros alcanzados en cuanto a satisfacción de necesidades básicas.

Gráfico 10  
**Población en pobreza coyuntural, 2014-2016**  
**(Porcentaje)**



Fuente: INEI - Encuesta Nacional de Hogares, 2007 - 2016.

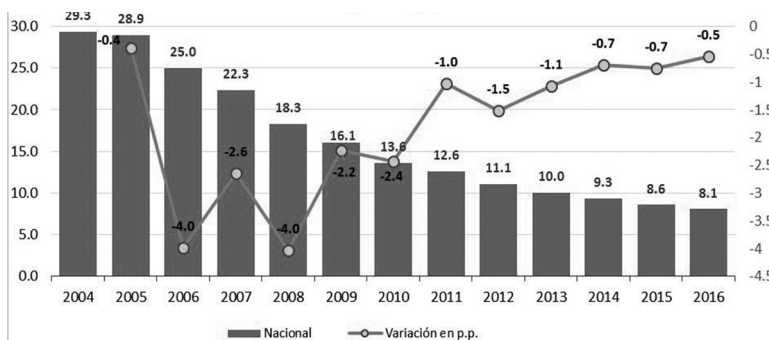
Durante el periodo 2004-2016 se observa una disminución en este tipo de población en 2.3 veces, dado que la población en pobreza coyuntural en el año 2004 era

el 29.4% y descendió en el año 2016 a 12.7%. En una menor proporción se aprecia la disminución entre el año 2007 y 2016, con un registro de 7.4 puntos porcentuales. Las diferencias nos muestran a una población en riesgo de caer en la pobreza crónica si las oportunidades laborales no les permiten recuperar su capacidad adquisitiva.

### Población en pobreza crónica

La pobreza crónica se refiere a la población que constituye el núcleo de la pobreza, no cuenta con un ingreso suficiente para un nivel mínimo de consumo, ni satisface sus necesidades más elementales. Es una población que se halla en una clara situación de exclusión social y, generalmente, posee más de una necesidad insatisfecha, por ejemplo: vive en condiciones prolongadas de privación, no puede acceder a los mínimos bienes y servicios, no ha logrado obtener una vivienda adecuada ni asegurar a todos sus miembros el acceso a la educación, a los servicios de salud y a oportunidades de empleo. Entre los años 2004 y 2016, la situación de estos pobladores se vio reducida en 21.3 puntos porcentuales y pasó de 29.3% (2004) a 8.1% (2016). Del mismo modo, este tipo de población se ha visto reducida en 2.8 veces, entre los años 2007 y 2016.

Gráfico 11  
**Población en pobreza coyuntural, 2014-2016**  
**(Porcentaje)**



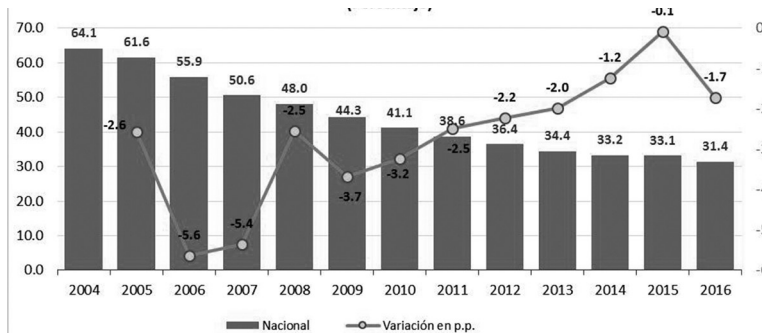
Fuente: Instituto Nacional de Estadística e Informática - Encuesta Nacional de Hogares, 2007 - 2016.

### Población en situación de pobreza integrada

Es aquella población que está ubicada en algunas de las categorías de pobreza propuestas mediante este método; es decir podría estar en condiciones de pobreza crónica, estructural o con carencias inerciales, o pobreza coyuntural o reciente. Ello significa que el porcentaje de la población en situación de pobreza integrada se complementa exactamente con el porcentaje de la población integrada socialmente.

De modo que las diferencias porcentuales entre los periodos 2004-2016 y 2007-2016 son de 32,7 y 19,2 puntos porcentuales, respectivamente. Asimismo, el año 2016 hubo una disminución de 1,7 puntos porcentuales y se pasa de 33,1 % (2015) a 31,4 % (2016).

Gráfico 12  
**Población en situación de pobreza integrada, 2014-2016**  
**(Porcentaje)**



Fuente: Instituto Nacional de Estadística e Informática - Encuesta Nacional de Hogares, 2007 - 2016.

## Análisis histórico de los factores determinantes del ingreso y de la pobreza

### Determinantes del ingreso de los hogares

Se han considerado de acuerdo con estudios realizados con anterioridad (Robles, 1997) como determinantes del ingreso de los miembros del hogar el sexo, la edad, el nivel educativo de la persona de 15 años y más. En el plano laboral afecta el ingreso de las personas la actividad económica a la que se dedica, la ocupación que tiene, el tamaño de la empresa en la que labora, entre las más importantes.

### El factor personal y educativo

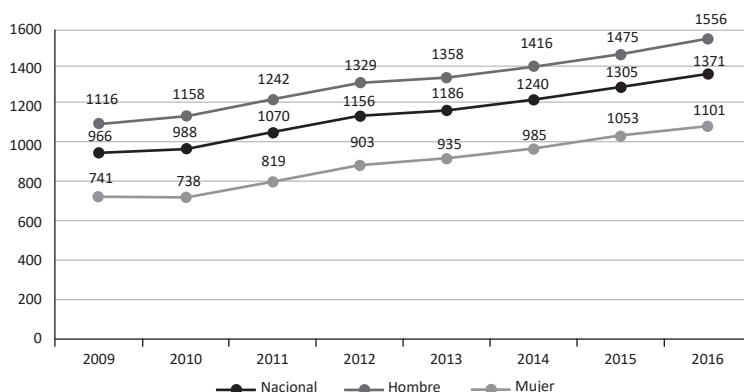
En este factor se han considerado las variables sexo, edad y nivel educativo.

#### Según sexo

Existe una persistente diferencia del ingreso en favor de la población masculina en desmedro de la femenina. Si se toma en consideración los ingresos promedio por trabajo de la población ocupada, según sexo, durante el periodo 2009-2016 los mayores ingresos promedio lo tienen los hombres cuya razón es de 1.5, aproximadamente, en cada uno de estos años. Es así, que en el año 2016 los ingresos promedio de los

hombres y las mujeres fueron de 1556 soles y 1101 soles, respectivamente, y alcanzó una variación promedio respecto del año anterior de 5,5 % y 4,6 %, respectivamente.

**Gráfico 13**  
**Ingreso promedio por trabajo de la población ocupada, según sexo, 2009-2016**  
**(soles corrientes)**

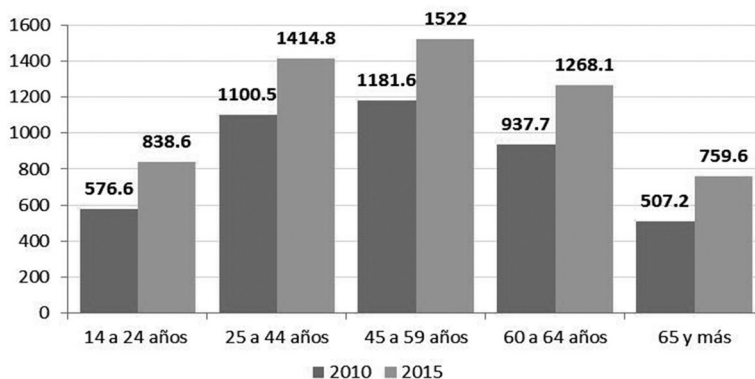


Fuente: INEI - Encuesta Nacional de Hogares, 2009 - 2016.

### Según grupos de edad

Igualmente se advierten diferentes ingresos asociados con la edad. Hasta los 25 años los jóvenes están en formación académica y entre los 25 y 35 años complementan su formación con estudios de posgrado. Entre los 35 y 45 tiene un plus valor la experiencia que se ha ganado en los sucesivos puestos.

**Gráfico 14**  
**Ingreso promedio mensual proveniente del trabajo, según grupos de edad,**  
**2010 y 2015 (soles corrientes)**

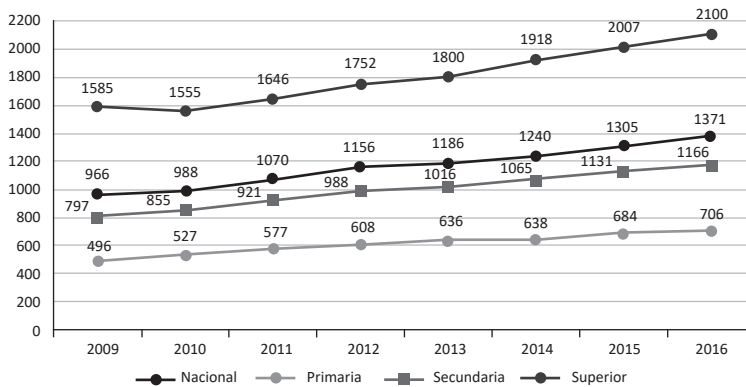


Fuente: INEI - Encuesta Nacional de Hogares, 2007 - 2016.

### Según nivel de estudios

Durante el periodo 2009-2016, los ingresos promedio de la población ocupada, según nivel de estudios, ponen en evidencia la relación existente entre nivel educativo y nivel de ingreso. Es así que la persona más preparada podrá generar productos o servicios con un mayor valor agregado, los cuales repercutirán en el valor final en el mercado, que a su vez hará que los ingresos percibidos por este grupo de personas se incrementen.

Gráfico 15  
**Ingreso promedio por trabajo de la población ocupada, según nivel de estudio, 2009-2016 (soles corrientes)**



Fuente: INEI - Encuesta Nacional de Hogares, 2009-2016.

De este modo se observa que en el año 2016 el ingreso percibido por la población con nivel superior de estudios es de 2 100 soles, y resalta una diferencia promedio de 934 soles con aquellos que tienen secundaria y una 1 394 soles con aquella población que cuenta con nivel de educación primaria.

### El factor laboral relacionado con la calidad del empleo

En esta categoría se considera según rama de actividad, categoría ocupacional y tamaño del centro laboral.

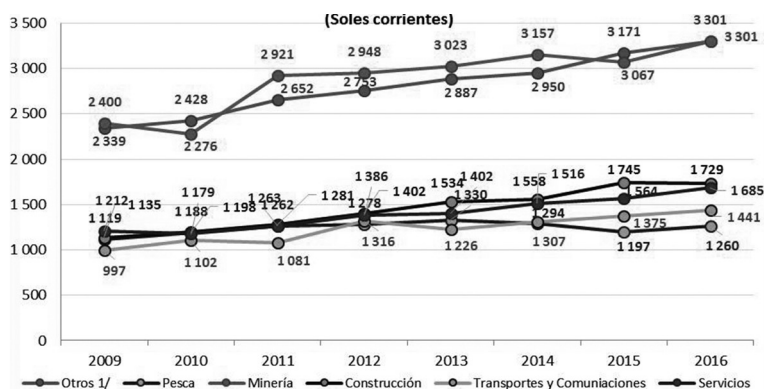
#### Según rama de actividad

La actividad que agrupa a la población con mayores ingresos promedio por su trabajo en el periodo 2009-2016 es la actividad minera, la cual alcanzó en el 2016 un promedio de 3 301 soles. Las diferencias más resaltantes en comparación con otras actividades son: 2 639 soles con la agricultura, 2 065 soles con el comercio, 2 041 soles con la pesca y 1 898 soles con la manufactura.



Gráfico 16

**Ingreso promedio por trabajo de la población ocupada, según rama de actividad, 2009-2016 (soles corrientes)**



Otros: incluye agricultura, comercio y manufactura  
 Fuente: INEI - Encuesta Nacional de Hogares, 2009 - 2016.

**Según categoría ocupacional**

Si los ingresos se ordenan de mayor a menor, los mayores le corresponden al empleador o patrono (por ser el propietario de la empresa), luego a los empleados que están mejor calificados y finalmente a los obreros que realizan actividades manuales que requieren baja calificación. En el caso de los independientes, la mayoría es subempleada debido a que tienen un autoempleo como alternativa al desempleo y sus ingresos incluso son superados ligeramente por los trabajadores del hogar.

**Según tamaño del centro laboral**

Casi la mayoría de las medianas empresas son formales y otorgan mejores ingresos, pero no superan a las empresas más grandes porque estas utilizan más tecnología y economías de escala.

En el año 2015 el nivel de ingreso de los trabajadores de empresas de 51 y más trabajadores fue de 2 162 soles; en las medianas empresas se pagaba en promedio 1 637 soles, mientras que en las empresas más pequeñas se pagaba en promedio una remuneración ligeramente más alta que la remuneración mínima vital. Aunque la variación promedio anual de los ingresos de los trabajadores de empresas de 1 a 10 trabajadores es más alto, en términos absolutos es menor (Anexo 6).

**Determinantes de la pobreza**

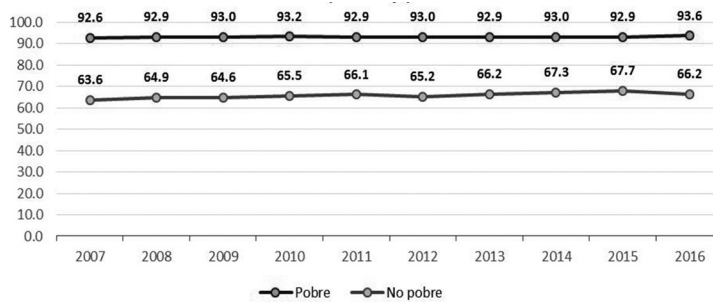
A continuación, se analiza la evolución de las variables que tienen una característica estructural e impactan en la incidencia de la pobreza.

### Nivel educativo alcanzado

La población en situación de pobreza muestra en promedio un menor nivel de educación, respecto de la no pobre. Más del 92 % de la población de 15 años y más, que alcanzó los niveles de primaria o secundaria, se encuentra en las categorías de los pobres. En cambio, en la población no pobre, más del 33 % de personas que tienen 15 años y más tienen formación superior.

Gráfico 17

**Población de 15 y más años de edad con nivel alcanzado de primaria y secundaria, según condición de pobreza, 2007-2016 (porcentaje)**



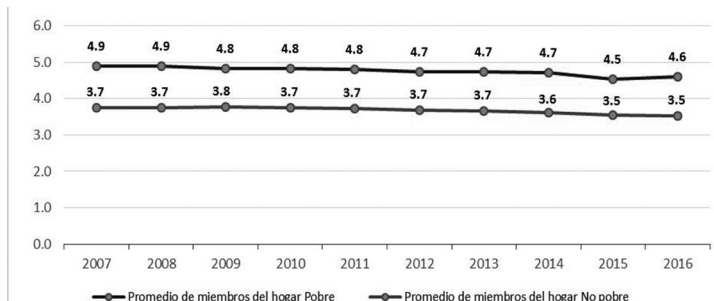
Fuente: INEI - Encuesta Nacional de Hogares, 2007-2016.

### Número promedio de los miembros del hogar

El tamaño del hogar es otro de los factores que genera más riesgo de pobreza. Entre los años 2007 y 2016, los hogares pobres concentraban un mayor número de miembros que los hogares no pobres, con una fluctuación promedio entre 4,5 y 4,9 miembros los hogares pobres en ese periodo. Para el caso de los hogares no pobres el número de miembros promedio fluctúa entre 3,5 y 3,8.

Gráfico 18

**Promedio de miembros del hogar, según condición de pobreza, 2007-2016 (número de personas)**



Fuente: INEI - Encuesta Nacional de Hogares, 2007-2016.

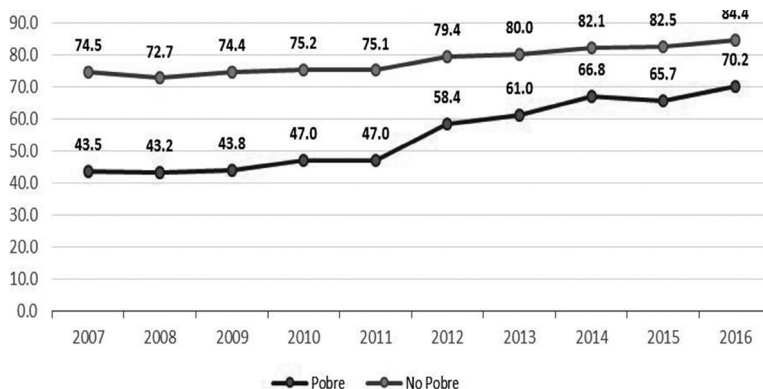
### Abastecimiento de agua por red pública dentro de la vivienda

Durante el periodo 2007-2016 se logró incrementar en forma sustancial el abastecimiento de agua por medio de red pública en las viviendas en condición de pobreza, lo que generó un incremento de 26,7 puntos porcentuales en dicho periodo, ya que el porcentaje de viviendas con agua por medio de red pública se elevó de 43,5 % a 70,2 %.

Por el lado de los hogares no pobres, dicho abastecimiento también presentó un leve incremento de 9,9 puntos porcentuales de 2007 a 2016. Cabe precisar que fue a partir del año 2013 en que se notó una ligera disminución de la brecha entre hogares y que la diferencia entre ambos tipos de hogares fue de 14,2 puntos porcentuales en el año 2016.

Gráfico 19

#### Abastecimiento de agua por red pública dentro de la vivienda, según condición de pobreza, 2007-2016 (porcentaje)



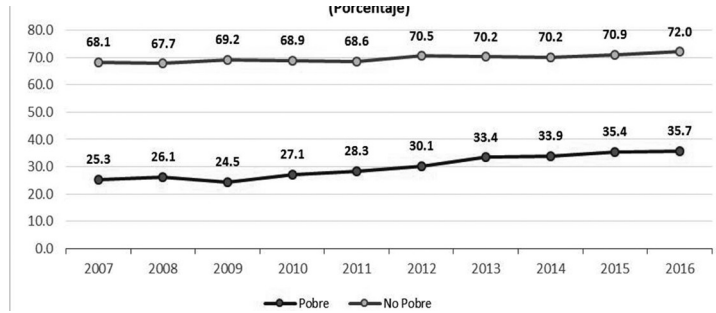
Fuente: INEI - Encuesta Nacional de Hogares, 2007-2016.

### Servicios higiénicos por red pública dentro de la vivienda

En los hogares en condición de pobreza, este tipo de servicio ha tenido un ligero incremento durante el periodo 2007-2016, pues alcanzó los 10,4 puntos porcentuales. Si se toma a los hogares no pobres se puede observar un incremento en el mismo periodo de tiempo de 3,9 puntos porcentuales, dado que el porcentaje de viviendas con este tipo de servicio subió de 68,1 % en 2007 a 72,0 % en 2016.

Las brechas entre los hogares no pobres y los de situación de pobreza también favorecen a los no pobres en lo que respecta a disponer de este servicio básico y es desde el año 2012 en que estas brechas se han venido acortando y han llegado a diferenciarse en 36,3 puntos porcentuales en el año 2016.

Gráfico 20  
**Servicios higiénicos por red pública, según condición de pobreza, 2007-2016 (porcentaje)**

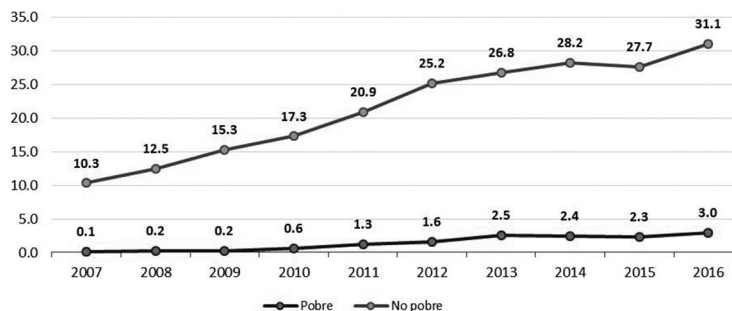


Fuente: INEI - Encuesta Nacional de Hogares, 2007-2016.

### Hogares con acceso a internet

Uno de los principales instrumentos para la lucha contra la pobreza y la desigualdad social es la inclusión digital. El incremento de estas brechas de acceso a internet es progresivo y llegó a presentar en el año 2016 una diferencia de 28,1 puntos porcentuales en relación con el año anterior. Este fenómeno se puede explicar en alguna medida debido a que las fases de acceso a nuevas tecnologías comienzan con los riesgos para que se extiendan a los pobres, que luego de un tiempo más o menos extenso (el ciclo muchas veces no se completa), conlleva a que la introducción de las nuevas tecnologías esenciales para el desarrollo haga que se incremente la desigualdad. En general, la infraestructura y las condiciones de vida que tienen los hogares no pobres, permite una mayor velocidad en el acceso al internet respecto a los hogares en condición de pobreza.

Gráfico 21  
**Hogares con acceso a internet, según condición de pobreza, 2007-2016 (porcentaje)**



Fuente: INEI - Encuesta Nacional de Hogares, 2007-2016.

## Resultados

### Determinantes de la pobreza 2016 en modelo

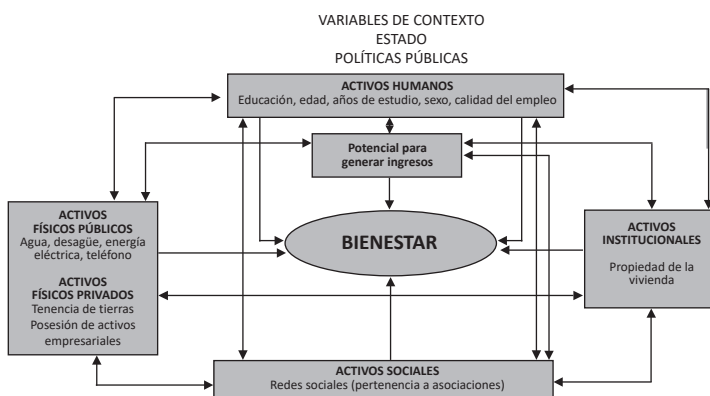
Hay variables que explican la variación del ingreso y de la pobreza. Esta relación paralela se explica en una línea de dependencia, en función a la teoría económica del consumidor. En efecto, el gasto en consumo está en función del nivel de ingreso a partir de una propensión marginal a consumir. Asimismo, el gasto es la variable utilizada para calificar si un hogar es pobre sobre la base de una línea de pobreza. Por tanto, el modelo de variables que explican el ingreso tiene algunos componentes similares al modelo de pobreza. En el año 2009, precisamente, se realizó una investigación sobre los factores que determinaban la pobreza.

El análisis por regresión econométrica permitió cuantificar el impacto de cada una de las características demográficas y socioeconómicas en la probabilidad de ser pobre. Los factores asociados a una menor probabilidad de ser pobre, un mayor número de perceptores de ingreso y tener acceso a servicios básicos como agua potable y electricidad están asociados a un gasto más alto (y a menor pobreza). Por el contrario, vivir en un hogar con mayor tamaño, cuyo jefe de hogar es joven o mujer, o tiene poca educación o mala calidad del empleo aumenta la probabilidad de ser pobre. (Banco Mundial, 2009).

Nuestro estudio se aborda desde un enfoque estructural, que incorpora dentro de sus modelos explicativos los activos de los pobres y la estructura de oportunidades provenientes del mercado, del Estado y de la sociedad.

Gráfico 22

#### Esquemas de activos - estructura de oportunidades



Fuente: Adaptado parcialmente del documento *Microcrédito y Finanzas* de Alberto Martínez

Para fines de comparación, la determinación de los factores que explican la pobreza ha utilizado las variables y el modelo de regresión logística del documento *Determi-*

*nantes de la Pobreza 2009* del INEI, que se ha actualizado con la base de datos de la Encuesta Nacional de Hogares del año 2016.

Las variables para el 2016 están estructuradas de la siguiente manera:

Tabla 3  
**Tipo de variable y recodificación**

Tipo de variable		Recodificación
<b>DEPENDIENTE</b>		
Y	Condición de pobreza	CONDICION_POBREZA
<b>INDEPENDIENTES</b>		
<b>Características del hogar</b>		
X1	Tamaño del hogar	MIEPERHO_
X2	Número de perceptores de ingresos	PERCEPHO
X3	Año de educación promedio del resto del hogar (Mayores de 14 años)	ANO_EDUCACION_RESTO_HOGA
<b>Características individuales</b>		
X4	Sexo del jefe del hogar	SEXO_JEFE
X5	Edad del jefe del hogar	CALIDAD_EMPLEO_JEFE
X6	Años de estudios del jefe del hogar	DISPONIBILIDAD_SERVICIOS_HIGIENICOS
X7	Calidad del empleo del jefe del hogar	DISPONIBILIDAD_ENERGIA_ELECTRICA
X12	Disponibilidad de servicios higiénicos	TENENCIA_TELEF
X13	Disponibilidad de energía eléctrica	PERTENENCIA_ASOCIACIONE
X14	Tenencia de teléfono	AÑOESTUDIOJEFE
X16	Pertenencia a asociaciones	EDADDEJEFE

Para el año 2009 las variables y categorías consideradas para el respectivo análisis de los determinantes de la pobreza, fueron:

Tabla 4  
**Perú: variables consideradas en el análisis de los determinantes de la pobreza, 2009**

Variables		Categorías
Condición de pobreza	Y	1 - Pobre 0 - No pobre
<b>Características del hogar</b>		
Tamaño del hogar	X1	[1a 17 miembros]
Número de perceptores de ingresos	X2	[1a 9 perceptores]
Años de educación promedio del resto del hogar (mayores de 14 años)	X3	[0 a 18 años]

<b>Características individuales</b>		
Sexo del jefe	X4	1 - Mujer 0 - Hombre
Edad del jefe	X5	(16 - 98 años]
Años de estudio del jefe	X6	1 0 a 18 años]
Calidad del empleo del jefe	X7	1 - Mala calidad del empleo 0 - Regular o buena calidad del empleo
Número de ocupaciones del jefe de hogar	X8	1 - Con 2 empleos 0 - Con 1 empleo
<b>Capital físico privado</b>		
Tenencia de tierras (algún miembro del hogar es propietario de la tierra)	X9	1 - Propia 0 - Caso contrario
Posesión de activos empresariales	X10	1 - Trabajo, trabajo y hogar 0 - Hogar o ninguno
<b>Capital físico público (Servicios básicos de la vivienda)</b>		
Disponibilidad de agua	X11	1 - Dispone 0 - No dispone
Disponibilidad de servicios higiénicos	X12	1 - Cuenta con servicios higiénicos 0 - No cuenta con servicios higiénicos
Disponibilidad de energía eléctrica	X13	1 - Cuenta con alumbrado eléctrico 0 - No cuenta con alumbrado eléctrico
Tenencia de teléfono	X14	1 - Cuenta al menos con telefonía fija o celular 0 - No cuenta con telefonía fija o celular
<b>Capital institucional</b>		
Propiedad de la vivienda	X15	1 - La vivienda tiene título de propiedad 0 - Está en trámite o no tiene título de propiedad
<b>Capital social</b>		
Pertenencia a asociaciones	X16	1 - Cuenta con activos sociales 0 - No cuenta con activos sociales

Fuente: INEI-ENAH0 2009. Elaboración: CIDE

## Resultados de la estimación del modelo Logit a nivel nacional

### Para el modelo que explique la pobreza

Con la aplicación de las mismas categorías se presenta los resultados del modelo del logit binario, en el que se observa, en términos estadísticos, si las variables son significativamente explicativas (Sig.), así como el análisis del Odd Ratio a través de la lectura del Exp (B), el cual nos muestra la intensidad del impacto. De la columna de los "B" extraemos el signo para indicar si la relación es directa cuando es positiva;

o existe una relación inversa cuando los resultados son negativos. Para estimar los resultados del modelo logit a nivel nacional se consideró: condición de pobreza =1 y condición no pobre = 0

Es importante mencionar que las variables relacionadas con el capital físico privado, tenencia de tierras y posesión de activos empresariales, no entraron en el modelo porque no se encontró en las bases de datos del 2016, las que fueran similares a las utilizadas en el año 2009. Asimismo, la variable Número de ocupaciones del jefe del hogar no era estadísticamente significativa como para tener impacto en la condición de pobreza por lo que no se le consideró en los resultados finales.

Tabla 5  
Perú: Modelo de regresión logística de los determinantes de la pobreza, 2016

	Variables en la ecuación							
	B	Error estándar	Wald	gl	Sig.	Exp(B)	95% CI para EXP(B) Inferior Superior	
Paso 1ª MIEPERHO	,494	,013	1526,795	1	,000	1,640	1,599	1,681
PERCEPHO	-,417	,023	327,785	1	,000	,659	,630	,689
Año_educacion_resto_hogar_mean	-,315	,016	372,760	1	,000	,730	,707	,754
SexoJefe	-,334	,052	42,032		,000	,716	,647	,792
Calidad_empleo_jefe	1,703	,077	486,050	1	,000	5,491	4,720	6,389
Disponibilidad_servicios_higiénicos	-,392	,053	55,451	1	,000	,676	,610	,749
Disponibilidad_energía_eléctrica	-,252	,055	21,099	1	,000	,777	,698	,865
tenencia_telef	-,644	,055	135,629	1	,000	,525	,471	,585
Pertenencia asociaciones	,357	,039	82,743	1	,000	1,429	1,323	1,544
AñoEstudioJefe	,022	,013	3,156	1	,076	1,022	,998	1,048
EdadDelJefe	,004	,001	7,176	1	,007	1,004	1,001	1,006
Constante	-1,358	,154	77,526	1	,000	,257		

a. Variables especificadas en el paso 1: MIEPERHO. PERCEPHO. Año\_educacion\_resto\_hogar\_mean. SexoJefe, Calidad\_empleoJefe. Disponibilidad\_servicios\_higiénicos. Disponibilidad\_energía\_eléctrica. tenencia\_tetef. Pertenencia\_asociaciones. AñoEstudioJefe. EdadDeJefe.

Este análisis consiste en analizar si aumenta o disminuye la probabilidad de ser pobre o no pobres, mediante la relación de cada una de las características y manteniendo las otras “ceteris paribus”, según los resultados del modelo logit.

Con excepción de la variable incluida “Año Estudio Jefe” que explica significativamente al 92% de confianza; el resto de variables tienen un nivel de confianza en la explicación de la condición de pobreza superior al 98%.



## **Variables que aumentan la probabilidad de ser pobre**

El tamaño del hogar (MIEPERHO). Cuando aumenta de miembros del hogar. Es decir, si el tamaño del hogar se eleva en un miembro, sube la probabilidad de que el hogar sea pobre en 64%. En efecto, ante a un mismo nivel de ingresos del hogar, cuando el tamaño del hogar aumenta, habrá un aumento de la demanda del consumo del hogar, sin embargo se reduce el consumo per cápita.

La calidad del empleo del jefe del hogar. Cuando el jefe del hogar cuenta con un empleo de mala calidad. En cuanto a los factores vinculados con el mercado laboral, en el Perú es notable el impacto de la mala calidad en el empleo sobre la probabilidad de ser pobre; es así que cuando el jefe del hogar se encuentra en situación de subempleo se eleva la probabilidad de que el hogar sea pobre en 491%, con respecto al jefe que cuenta con un empleo de buena calidad, operativizado con el adecuadamente empleado. Esta constatación pone en evidencia los vínculos entre el empleo de mala calidad que incide en las bajas retribuciones del jefe del hogar.

Años de estudio del jefe del hogar. Adicionalmente a la menor significancia estadística de esta variable, no hay un gran impacto en el aumento o reducción con respecto al aumento de un año de estudio en la educación del jefe del hogar en el año 2016, en relación con la probabilidad de ser pobre. Ello se puede explicar al mayor promedio de años de estudios que tienen los jefes del hogar respecto a los otros miembros: un año de instrucción superior. Se puede interpretar de la siguiente manera: puede tener un o dos años de instrucción superior en promedio, un año más no afecta su condición de pobreza; lo cual nos hace pensar que depende más de la educación de los otros miembros del hogar para acceder a nuevas oportunidades de desarrollo porque el nivel promedio de ellos es más bajo.

Pertenencia a asociaciones. La pertenencia del jefe del hogar a algún club o asociación deportivos, agrupación o partido político, club cultural (de danza o música, u otra), etc., eleva la probabilidad de que el hogar sea pobre en 42,9 %, con respecto al hogar que no pertenece a alguna asociación.

Edad el jefe de hogar. Se advierte una ligera relación directa entre la edad del jefe del hogar y la probabilidad de ser pobre, pues cuando aumenta la edad del jefe del hogar, la probabilidad de que el hogar sea pobre aumenta en 0,4 %; porcentaje casi insignificante que sin embargo refleja que el mercado laboral poco a poco prefiere a los jefes de hogar más jóvenes para ciertas tareas debido a estos les generan menores costos laborales, a diferencia de los jefes de hogar mayores.

## **Variables que disminuyen la probabilidad de ser pobre**

Se confirma la relación inversa entre las siguientes variables y la condición de pobreza del hogar:

El número de perceptores (PERCEPHO). Un aumento del número de perceptores del hogar aumenta la capacidad de generación de más ingreso en el hogar y disminuye la probabilidad de que el hogar sea pobre en un 34,1 %. De lo cual se deduce que lo que incide realmente en la reducción de la probabilidad de que un hogar sea pobre en el Perú, no es que el jefe de hogar trabaje o cuente con dos empleos o más sino que en el hogar existan varios aportantes de ingresos.

Años de educación promedio del resto de miembros del hogar. Se ha cuantificado que si aumenta en un año el promedio de años de estudios del resto de los miembros del hogar, disminuye la probabilidad de que el hogar sea pobre en un 27 %. Lo cual es razonable, dado que los adultos dentro del hogar tienen en promedio años de estudios menores al jefe del hogar. Mayores años de estudio generan más posibilidades de acceso a mejores empleos; por ello, el Banco Mundial considera que la acumulación del capital humano permite mejorar directamente el bienestar y aumenta la habilidad de los más pobres para aprovechar las oportunidades de ingresos ofrecidos por el mercado.

Provisión de energía eléctrica a los hogares. Dentro del acceso a los servicios básicos de la vivienda, este tiene un efecto claramente positivo sobre la situación del hogar y disminuye la probabilidad de ser pobre en un 22,3 %, en razón de las oportunidades que el uso productivo de la energía ofrece y de potencializa las perspectivas de aumento del ingreso familiar.

Disponibilidad de servicios higiénicos de la vivienda. Tiene un efecto positivo en la disminución de la probabilidad de ser pobre en un 32,4 % por las oportunidades que genera su uso productivo. La carencia de estos servicios refleja la necesidad de mejorar las condiciones de vida de la población y de evitar que contraiga enfermedades.

Disponibilidad de teléfono fijo o celular. Tiene un efecto positivo relevante en la disminución de la probabilidad de ser pobre. Nuevamente, el signo negativo del coeficiente asociado a dicha variable refleja que cuando se dispone de telefonía fija o de celular, existe menos probabilidad de ser pobre, pues esta disminuye en 47,5 %. Lo que se explica igualmente por las oportunidades de generación de ingresos que se derivan del uso productivo de estos activos.

## **Resultados de la estimación del modelo Logit a nivel nacional para explicar la pobreza extrema**

Este análisis consiste en determinar las variables que aumentan o reducen la probabilidad de los hogares de estar en situación de pobreza extrema. Hay que tomar en consideración que las mismas variables que explicaron el aumento de la situación de pobreza, se han reconsiderado para determinar el impacto en el aumento de la probabilidad de ser pobre extremo.

A continuación los resultados del modelo logit.

Tabla 6

**Perú: Modelo de regresión logística de los determinantes de la pobreza extrema, 2016**

	Variables en la ecuación							
	B	Error estándar	Wald	gl	Sig.	Exp(B)	95% CI para EXP(B) Inferior Superior	
Paso 1 <sup>a</sup> MIEPERHO	,405	,020	429,648	1	,000	1,499	1,443	1,558
PERCEPHO	-,365	,044	69,323	1	,000	,000	,637	,757
Año_educacion_resto_hogar_mean	-,397	,033	140,293	1	,000	,673	,630	,718
SexoJefe	-,370	,116	10,277	1	,001	,691	,551	,866
Calidad_empleo_jefe	2,643	,359	54,072	1	,000	14,052	6,948	28,423
Disponibilidad_servicios_higiénicos	-,304	,086	12,523	1	,000	,738	,623	,873
Disponibilidad_energía_eléctrica	-,408	,086	22,541	1	,000	,665	,562	,787
tenencia_telef	-,607	,087	48,481	1	,000	,545	,459	,646
Pertenencia_asociaciones	,520	,090	33,550	1	,000	1,682	1,411	2,006
AñoEstudioJefe	,040	,025	2,594	1	,107	1,041	,991	1,092
EdadDelJefe	,008	,003	7,934	1	,005	1,008	1,002	1,013
Constante	-4,173	,436	91,827	1	,000	,015		

a. Variables especificadas en el paso 1: MIEPERHO. PERCEPHO. Año\_educacion\_resto\_hogar\_mean. SexoJefe, Calidad\_empleoJefe. Disponibilidad\_servicios\_higiénicos. Disponibilidad\_energía\_eléctrica. tenencia\_tetef. Pertenencia\_asociaciones. AñoEstudioJefe. EdadDelJefe.

Los resultados del modelo logit se presentan en la tabla 6, en la cual cabe precisar que la categoría de la variable incluida “AñoEstudioJefe” no es significativa en el agregado nacional, pero se la ha mantenido debido a que sí lo es a nivel de área urbana o rural. En ese sentido, las interpretaciones se harán sobre las variables significativas a un nivel del cinco por ciento, lo que significa que se tiene el 95 por ciento de confianza para generalizar sin equivocaciones.

### **Variables que aumentan la probabilidad de ser pobre extremo**

El tamaño del hogar (MIEPERHO). Tal como se esperaba, aumenta la probabilidad de que el hogar sea pobre extremo si el número de miembros se incrementa; y existe la probabilidad de que esta se eleve en 49.9% con respecto al hogar cuyo tamaño es más pequeño. Es por ello, que frente a un mismo nivel de ingresos, cuando el tamaño del hogar es grande, más miembros participarán de la distribución del mismo, viendo sus opciones más reducidas que cuando el tamaño de hogar es más pequeño.

La calidad del empleo del jefe del hogar. El impacto de la calidad del empleo (Calidad\_empleo\_jefe) del jefe del hogar sobre la probabilidad de ser pobre extremo en el Perú, es considerable; de modo que cuando el jefe del hogar cuenta con un empleo de mala calidad (empleo informal), se eleva la probabilidad de que el hogar sea pobre en 1305%, con respecto al jefe de hogar que cuenta con un empleo de regular o buena calidad (empleo formal).

Años de estudio del jefe del hogar. Al igual que en el modelo de pobreza, esta variable no presenta un gran impacto en el aumento o reducción de la pobreza extrema, respecto al aumento de un año de estudio en la educación del jefe del hogar, lo cual nos hace pensar que depende más de la educación del resto de miembros del hogar, para acceder a nuevas oportunidades de desarrollo.

No pertenencia a asociaciones. Cuando el jefe de hogar no pertenece a ninguna asociación, llámese clubes y asociaciones deportivas, agrupación o partido político, clubes culturales (danza, música, etc.), entre otras formas de asociación, la probabilidad de ser pobre extremo se eleva en 68.2%, con respecto al jefe de hogar que está vinculado a alguna asociación.

## **Variables que disminuyen la probabilidad de ser pobre extremo**

Las siguientes variables analizadas, según los resultados del modelo aplicado, determinan una relación inversa entre estas y la pobreza extrema en el hogar.

Número de perceptores (PERCEPHO). El incremento del número de perceptores en el hogar genera mayores posibilidades de producir más ingresos y ello repercute en la disminución de la probabilidad en un 30.6% de que el hogar sea pobre extremo.

Años de educación promedio del resto del hogar. Se puede determinar que el incremento en un año el promedio de estudio del resto de miembros del hogar, genera una disminución en la probabilidad de que el hogar sea pobre extremo, en un 32.7%.

Dentro del acceso a los servicios básicos de la vivienda, tenemos las siguientes variables:

Provisión de energía eléctrica. El contar con este tipo de servicio básico en los hogares tiene un efecto claramente positivo sobre la situación del hogar y disminuye la probabilidad de ser pobre extremo en 33.5%.

Disponibilidad de servicios higiénicos de la vivienda tiene un efecto positivo en la disminución de probabilidad de ser pobre extremo, en un 26.2%.

Disponibilidad de teléfono fijo o celular. La tenencia de este tipo de tecnologías de la comunicación genera una menor probabilidad de ser pobre, situación que disminuye en 45,5%. Lo que se explica igualmente por las oportunidades de generación de ingresos, derivadas del uso productivo de estos activos.

**Tabla 7**  
**Anexo. Modelo logit en Eviews**

Dependent Variable: CONDICION\_POBREZA  
 Method ML - Binary Logit (Quadratic hili dimbing)  
 Date 08/29/17 Time: 0957  
 Sample: 1 35785  
 Inlucaded observations 27953  
 Convergence achieved after 6 iterations  
 Covartance matrix computed using second derivativos

Variable	Coefficient	Std Error	z-Statistic	Prob.
MIEPERHO	0.494434	0.012654	39.07423	0.0000
PERCEPHO	-0.417064	0.023036	-18.10485	0.0000
ANO_EDUCACION_RESTO_HOGA	-0.314731	0.016301	-19.30700	0.0000
SEXOJEFE	-0.334214	0.051551	-6.483192	0.0000
EDAD16 98	0.003704	0.001383	2.678794	0.0074
CALIDAD_EMPLEO_JEFE	1.703145	0.077252	22.04654	0.0000
DISPONIBILIDAD_SERVICIOS	-0.391802	0.052615	-7.446540	0.0000
DISPONIBILIDAD_ENERGIA_E	-0.252104	0.054884	-4.593376	0.0000
TENENCIA_TELEF	-0.643933	0.055292	-11.64601	0.0000
PERTENENCIA_ASOCIACIONES	0.357129	0.039261	9.096343	0.0000
ANOESTUDIOJEFE	0.022245	0.012521	1.776596	0.0756
C	-1.357969	0.154229	-8.804872	0.0000
McFadden R-squared	0.249230	Mean dependentvar		0.182950
S.D. dependentvar	0.386632	S.E. of regression		0.335724
Akaike info criterion	0.715347	Sum squared restd		3149.244
Schwarz criterion	0.718884	Log likelihood		-9986.047
Hannan-Quinn criter.	0.716486	Deviance		19972.09
Restr. deviance	26602.16	Restr. log likelihood		-13301.08
LR statistic	6630.066	Avg. log likelihood		-0.357244
Prob(LR statistic)	0.000000			
Obs with Dep=0	22839	Total obs	27953	
Obs with Dep=1	5114			

## Análisis comparativo de resultados 2009 y 2016

En la tabla 8 se presenta un comparativo de los resultados del modelo de regresión logística aplicados a los años 2009 y 2016:

Tabla 8  
Modelo determinante de la pobreza en el Perú, 2009 y 2016

Variables		Odds	Ratio
		2009	2016
MI EPERHO	X1	1.627	1.640
PERCEPHO	X2	0.674	0.659
Año_educación_resto_hogar_mean	X3	0.845	0.730
Sexo_Jefe	X4	1.222	0.716
EdadDel Jefe	X5	0.978	1.004
AñoEstudioJefe	X6	0.932	1.022
Calidad_empleo_jefe	X7	2.928	5.491
Disponibilidad_servicios_higienicos	X12	0.886	0.676
Disponibilidad_energía_electrica	X13	0.851	0.777
Tenencia_te1ef	X14	0.524	0.525
Pertenencia_asociaciones	X16	0.560	1.429

*Nota: 2009 elaborado por CIDE - INEI y 2016 elaboración propia  
Fuente: INEI - ENAHO, 2009 y 2016*

**(X1).** Si se toma en cuenta que el tamaño del hogar incrementa la probabilidad de que los hogares sean más pobres, se observa una mayor probabilidad de hogares pobres en el año 2016 (64.0%) que en el 2009 (62.7%). Dicha diferencia se debería al incremento del tamaño en los hogares en el 2016.

**(X2).** En el modelo trabajado se puede advertir una relación inversa entre el número de perceptores del hogar y el nivel de pobreza para los años 2009 y 2016. Para ambos años nos indica que a mayor número de perceptores, menores son las probabilidades de que el hogar sea pobre, estas son de 32.6% y 34.1% para los años 2009 y 2016, respectivamente.

**(X3).** En ambos años se puede observar una relación inversa entre los años de educación promedio de los otros miembros del hogar y la condición de pobreza del hogar. Se registra un mayor incremento en el promedio de estudios de estos miembros del hogar en el año 2016, lo cual genera una mayor disminución en la probabilidad de que los hogares sean pobres, dado que la diferencia entre el año 2009 y 2016 es de 11.46 puntos porcentuales.

**(X5).** Para el año 2009 el modelo indica que cuando la edad del jefe de hogar aumenta, la probabilidad de que el hogar sea pobre es de 2,2 %. Para el año 2016 esta relación de los jefes de hogar de menor edad respecto a los de mayor edad, establecía que ambos tenían la misma probabilidad de ser pobres o no serlo. La diferencia entre ambos años es de 2,63 puntos porcentuales.

**(X6).** Uno de los factores condicionantes fundamentales para el tipo de ocupación y por ende para el nivel de ingresos, es el nivel educativo alcanzado, por lo que a menor nivel de educación del jefe de hogar son mayores las probabilidades de ser pobre. De los modelos aplicados en los años 2009 y 2016, se desprende para el primero de ellos que a mayor número de años de estudio del jefe de hogar, la probabilidad de ser pobre disminuye en 6,8 %. Lo contrario se observa para el año 2016, donde a menores años de estudio del jefe de hogar, la probabilidad de ser pobre aumenta en 2,2 %.

**(X7).** Las probabilidades de incremento de que los hogares sean pobres, de acuerdo con el nivel de calidad de empleo del jefe de hogar es considerable. Para el modelo del año 2009, cuando el empleo del jefe de hogar es de mala calidad, la probabilidad de ser pobre se eleva en 192,8 %. Por otro lado, para el año 2016 esta probabilidad se eleva a 449.1% debido al fuerte incremento de la informalidad en todas las actividades económicas que se desarrollan en el país.

**(X12).** El acceso a servicios higiénicos en las viviendas, tanto para el año 2009 y como para el año 2016, ha tenido un efecto positivo en la disminución de la probabilidad de ser pobre; probabilidad que fue mayor en el año 2016, con 32,4 %, frente al 11.4% presentado en el 2009.

**(X13).** Al igual que el acceso a los servicios higiénicos, la disponibilidad de energía eléctrica en los hogares también tiene un efecto positivo en la disminución de las probabilidades de ser pobre. Este servicio subió de 14,9 % (2009) a 22,3 % (2016).

**(X14).** La mayor disponibilidad de telefonía fija o de celular en los hogares es una referencia de la menor probabilidad de ser pobre. Las cifras de los años 2009 y 2016 arrojan 47,6 % y 47,5 %, respectivamente.

**(X16).** De acuerdo a los modelos presentados para el año 2009, la pertenencia a alguna asociación disminuye la probabilidad de ser pobre, la cual fue en ese año 44,0 %. Para el año 2016, la no participación en alguna asociación permitió que la probabilidad a ser pobre se elevara en 42,9 %.

## **Análisis vertical**

Las variables que tuvieron mayor impacto en la probabilidad de ser pobre en el año 2016 fueron:

- **MIEPERHO (X1).** Si el tamaño del hogar es grande, la probabilidad de que el hogar sea pobre aumenta en 64%, con respecto al hogar más pequeño.

- **Calidad\_empleo\_jefe (X7).** Cuando el jefe del hogar cuenta con un empleo de mala calidad (empleo informal), la probabilidad de que el hogar sea pobre se eleva en 449,1% con respecto al jefe de hogar que cuenta con un empleo de regular o buena calidad (empleo formal).
- **Pertenencia\_asociaciones (X16).** Cuando el jefe del hogar no pertenece a ninguna asociación, la probabilidad de que el hogar sea pobre se eleva en 42.9%, con respecto al hogar que pertenece a alguna asociación.

### **Análisis horizontal**

Las variables de mayor variación en los años 2009 y 2016, en lo referente a la probabilidad de ser pobre, son:

- **Sexo\_jefe (X4).** Aquí cambia la relación de probabilidad de ser pobre, pues resulta en una elevación del 22.2% para el 2009, respecto a una disminución de 28.4% para el 2016.
- **Calidad\_empleo\_jefe (X7).** Resulta en un incremento muy significativo en la probabilidad de ser pobre, pues el incremento pasa de 192.8% en el año 2009 a un 449.1% para el año 2016.
- **Disponibilidad\_servicios\_higiénicos (X12).** Presenta una disminución significativa de 11.4% en el 2009 a 32.4% para el 2016, en la probabilidad de ser pobre.
- **Pertenencia\_asociaciones (X16).** Se da un cambio en la relación de probabilidad de ser pobre, que resulta en una disminución de 44% para el 2009, respecto a una elevación de 42.9% para el 2016.

Las variables que estadísticamente se mantuvieron iguales en los años 2009 y 2016 son las siguientes:

- **MIEPERHO (X1).** Se observa un incremento en la probabilidad de ser pobre, de 62.7% (2009) a 64% (2016).
- **PERCEPHO (X2).** Se dio una disminución en la probabilidad de ser pobre, de 32.6% a 34.1% entre el año 2009 y el año 2016, respectivamente.
- **Año\_educación\_resto\_hogar\_mean (X3).** Se observa una disminución en la probabilidad de ser pobre, de 15.5% (2009) a 27% (2016).
- **Edad\_del\_Jefe (X5):** Se da una disminución en la probabilidad de ser pobre, de 2.2% (2009) a 0.4% (2016).
- **Año\_Estudio\_Jefe (X6).** Presenta una disminución de 6.8% en 2009 y un aumento de 2.2% en la probabilidad de ser pobre en 2016.



- **Disponibilidad\_energía\_eléctrica (X13)**. Se muestra una disminución de 14.9% a 22.3%, en la probabilidad de ser pobre, para los años 2009 y 2016, respectivamente.
- **Tenencia\_teléfono (X14)**: Se registra un incremento mínimo en la probabilidad de ser pobre de 47.6% (2009) a 47.5% (2016).

## Conclusiones

1. El crecimiento de la economía peruana ha generado las condiciones para un crecimiento sostenido del ingreso y el gasto real en el periodo 2009-2016; con mayor velocidad en los primeros años. Luego, es el principal factor de soporte explicativo que genera las condiciones y favorece una redistribución del crecimiento económico, la reducción de la pobreza y pobreza extrema coyuntural o monetaria, medida a través de la capacidad de gasto del hogar con respecto a la línea de pobreza. Es importante señalar que, igualmente, se redujo la severidad y la brecha de pobreza
2. Asimismo se pudo apreciar que igualmente se redujeron las otras manifestaciones de la pobreza, como son: la estructural medida como la proporción de hogares con al menos una de las necesidades básicas insatisfechas, y que refleja el grado de acceso a los servicios básicos de la vivienda o educación, depende de la política del gasto social en infraestructura sanitaria, eléctrica y educativa. Igualmente disminuyó la pobreza crónica y la integrada. La primera referida a los hogares que tienen simultáneamente la pobreza coyuntural y estructural, la segunda referida a la suma de la proporción de hogares que tienen la pobreza coyuntural o la pobreza estructural.
3. Como un impacto adicional, se advierte una reducción de la desigual distribución del ingreso por cualquier método de concentración. Es así que Gini disminuye de 0.50 a 0.46; y el índice de concentración de Theil se reduce de 0.43 a 0.36. Por todo lo enunciado, se aprecia un aumento vigoroso de la clase media.
4. La educación y las características personales tienen un efecto positivo sobre el ingreso de los jefes del hogar que trabajan, siendo el factor fundamental de la determinación del ingreso. En efecto, el sexo, la edad los años de estudio, la ocupación, actividad económica, el tamaño de la empresa, están entre los principales factores que inciden en el ingreso del Hogar.
5. Tomando en consideración, los ingresos promedio por trabajo de la población ocupada, según sexo, durante el periodo 2009-2016, los mayores ingresos promedio lo mantienen los hombres cuya razón es de 1.5 respecto de las mujeres aproximadamente en cada uno de estos años.
6. Con respecto a los factores determinantes de la pobreza, el nivel educativo alcanzado aparece nuevamente como un factor preponderante, donde el crecimiento en

puntos porcentuales de la población con estudios primarios y secundarios, entre los años 2007-2016 es mayor en la población no pobre con un crecimiento de 2,7 puntos porcentuales, pasando de 63,6% (cifra del año 2007) a 66.2% en el 2016.

7. El tamaño del hogar, aumenta la probabilidad de que el hogar sea pobre, cuando cuenta con un número grande de miembros del hogar, se eleva la probabilidad de que el hogar sea pobre en 64%, con respecto al hogar cuyo tamaño es pequeño.
8. En cuanto a la calidad del empleo del jefe del hogar, cuando el jefe del hogar cuenta con un empleo de mala calidad, se eleva la probabilidad de que el hogar sea pobre en 449,1%, con respecto al jefe de hogar que cuenta con un empleo de regular o buena calidad (empleo formal).
9. A mayor número de perceptores, menores son las probabilidades de que el hogar sea pobre, siendo estas probabilidades de 32.6% y 34.1%, para los años 2009 y 2016, respectivamente.
10. La pertenencia a alguna asociación disminuye la probabilidad de ser pobre, disminuyendo la probabilidad en 44.0% (2009). Para el 2016, la no participación en alguna asociación, permitió que la probabilidad a ser pobre se eleve en 42.9%

## **Recomendaciones**

1. Las políticas de estado siempre deben priorizar promover e impulsar las condiciones para un crecimiento de la actividad económica el principal factor explicativo para la reducción de la pobreza en sus diferentes manifestaciones.
2. En segundo término, se debe priorizar la mejora en el nivel educativo, que genera mayor conocimiento para el desarrollo de las personas , así como mejores calificaciones para una mayor oportunidad laboral con la correspondiente mejora en los niveles de productividad e ingreso.
3. La pobreza estructural, medida como la proporción de hogares con al menos una de las necesidades básicas insatisfechas, y que refleja el acceso a los servicios básicos de la vivienda o educación, depende de la política del gasto social en infraestructura sanitaria, eléctrica y educativa.
4. Asimismo, en la aplicación de los programas sociales, con recursos escasos, se debe optimizar la efectividad a través de una buena focalización del gasto social. Ello acontece en el programa Juntos que tiene bien dirigida su accionar. El programa otorga transferencias condicionadas a la búsqueda del hogar del acceso a los servicios de salud y educación.
5. Finalmente, un elemento que favorece el acceso es el aumento de la infraestructura vial que se ha visto en otros trabajos de investigación.

## Referencias bibliográficas

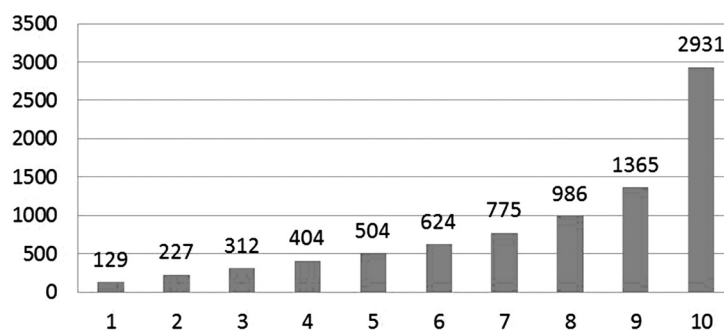
- Banco Mundial. (2009). *Una mirada a la evolución reciente de la pobreza en el Perú: Avances y desafíos*. Lima.
- Clodoaldo, C. (2004). *La evolución de la distribución del ingreso en el Perú*. Lima: Universidad Nacional Mayor de San Marcos.
- Di Virgilio, M., Otero, M. y Boniolo, P. (2010). *Pobreza y desigualdad en América Latina y el Caribe*. Buenos Aires: Consejo Latinoamericano de Ciencias Sociales.
- Filgueira, C. y Peri, C. (2004). *América Latina: Los rostros de la pobreza y sus causas determinantes*. Centro Latinoamericano y Caribeño de Demografía.
- Instituto Nacional de Estadística e Informática. (2017). *Evolución de la pobreza monetaria 2007-2016*. Lima: INEI.
- Instituto Nacional de Estadística e Informática. (2016). *Evolución de la pobreza monetaria 2009-2015*. Lima: INEI.
- Instituto Nacional de Estadística e Informática. (2015). *Evolución de la Pobreza Monetaria 2009-2014*. Lima: INEI.
- Instituto Nacional de Estadística e Informática. (2000). *Factores que determinan el ingreso de los hogares en el Perú*. Lima: INEI.
- Instituto Nacional de Estadística e Informática. (1998). *Pobreza, conceptos y métodos de medición*. Lima: INEI.
- López, R. (2010). *Pobreza y distribución del ingreso en América Latina*. CEPAL – AECID.
- Mendoza, W., Leyva, J. y Flor, J. (2011). *La distribución del ingreso en el Perú: 1980-2010*. Lima: Pontificia Universidad Católica del Perú.
- Muñoz, M. (2004). Determinantes del ingreso y del gasto corriente de los hogares. *Revista de Economía Institucional*.
- OCDE. (2016). *Estudio Multidimensional del Perú: Volumen 2, Análisis detallado y recomendaciones*. París: OCDE.
- Pérez, P. y Rodríguez, A. (2015). *El ejercicio de medir la pobreza en el Perú*. Lima: PRONABEC.
- Robles, M. (1997). *Factores determinantes de los ingresos del hogar*.
- Quispe, R. *Estudio de la evolución de la distribución del ingreso en el Perú, 1991-1999*. Lima: Universidad Nacional de Ingeniería.

Verdera, F. (2007). *La pobreza en el Perú: Un análisis de sus causas y de las políticas para enfrentarla*. Lima: Instituto de Estudios Peruanos.

## Anexos

Anexo 1

### Perú: ingreso per cápita mensual en hogares, 2016



Fuente: INEI-ENAHO 2016. Elaboración Propia

Anexo 2

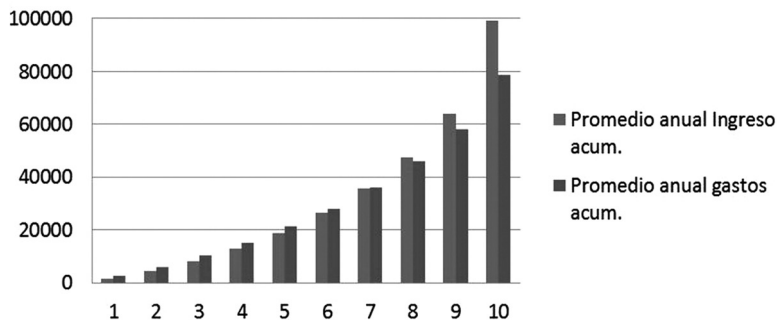
### Promedios anuales de ingreso, gasto, ahorro corriente, según deciles de ingreso a nivel nacional, 2016 (en soles)

Deciles de ingreso	Ingreso corriente	Gasto corriente	Ahorro corriente	Promedio de miembros	Promedio de perceptores
1	1 544.80	2 643.21	-1 098.41	4.49	1.81
2	4 263.54	6 077.12	-1 813.58	3.96	1.96
3	8 007.77	10 256.14	-2 248.37	3.81	1.99
4	12 859.99	15 278.81	-2 418.82	3.84	2.04
5	18 909.49	21 174.15	-2 264.66	3.88	2.18
6	26 394.51	28 068.88	-1 674.37	3.77	2.24
7	35 688.57	36 176.36	-487.80	3.62	2.27
8	47 524.45	45 880.59	1 643.85	3.38	2.28
9	63 900.64	58 177.17	5 723.46	3.13	2.24
10	99 075.92	78 634.74	20 441.18	2.6	1.96
Total	31 816.97	30 236.72	1 580.25	3.65	2.1

Fuente: INEI-ENAHO 2016. Elaboración propia.

Anexo 3

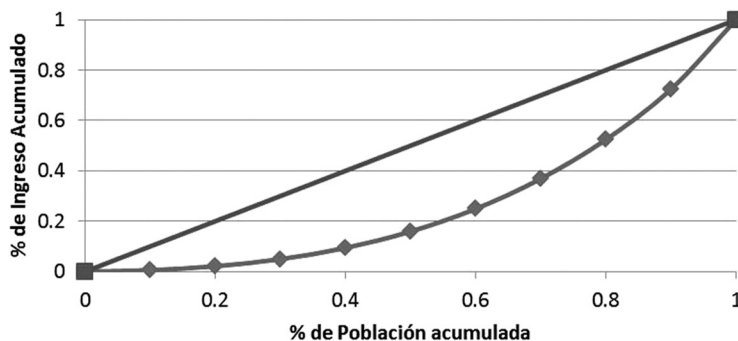
**Ingreso y gasto promedio per cápita anual, según deciles de ingreso por hogar a nivel nacional, 2016**



Fuente: INEI-ENAHO 2016. Elaboración propia.

Anexo 4

**Curva lorenz nivel nacional, 2016**



Fuente: INEI-ENAHO 2016. Elaboración propia.

Anexo 5

**Indicadores de desigualdad 2009, 2015 y 2016**

Año	Índice de Gini	Índice de concentración de Theil	Varianza del logaritmo	CV
2009	0.50409	0.429	0.877	1.1188
2015	0.45954	0.367		
2016	0.46073	0.361	0.74	1.0038

Fuente: INEI-ENAHO. Elaboración propia.

Anexo 6

**Perú: ingreso promedio mensual proveniente del trabajo, 2004-2015  
(Según tamaño de empresa, en soles corrientes)**

Tamaño de empresa	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	Tasa de crecimiento promedio anual (%) 2004-2015	Variación porcentual 2015/2014
Total	677.1	676.2	727.3	814.5	894.6	965.8	987.6	1 069.7	1 155.7	1 186.1	1 239.9	1 305.0	6.1	5.2
De 1 a 10	459.8	474.2	489.1	558.2	634.3	559.8	704.8	779.2	832.0	857.4	883.4	953.7	6.6	8.0
De 11 a 50	946.4	931.3	1 014.9	1 039.1	1 107.6	1 278.9	1 238.8	1 402.9	1 419.4	1 433.1	1 454.5	1 535.8	5.1	12.5
51 y más	1 395.6	1 326.8	1 448.4	1 589.1	1 662.4	1 801.5	1 796.9	1 817.4	1 952.0	2 017.1	2 132.4	2 162.2	4.1	1.4

Fuente: INEI-ENAHO, 2004-2015.

# Modelos ARIMA para el análisis de datos de sismos en el Perú en 2017

Carlos Risco Franco

## Resumen

Este proyecto busca explorar el fenómeno de los sismos mediante el uso de herramientas estadísticas, en este caso para caracterizar y buscar patrones de comportamiento de los mismos, que se han presentado en el Perú en el 2017; para ello se usan técnicas como el Cluster Análisis y Modelos de Series de Tiempo. Primero se ha identificado las zonas de mayor actividad sísmica y luego se ha hallado sus características principales, así como la interrelación entre magnitud y profundidad. Aparte de identificar las zonas sísmicas, se observa una pequeña tendencia a la disminución en la magnitud de los sismos en Lima en los últimos tres meses. Los datos se han ajustado a un modelo de series de tiempo ARIMA (1,1,0), el cual ha resultado significativo, tanto a nivel nacional, como para Lima-Ica y Arequipa. Se ha utilizado datos del Instituto Geofísico del Perú en su página web. La zona de mayor actividad es la zona de Arequipa, seguida de la de Lima.

**Palabras clave:** sismo, Cluster Analysis, Modelos de Series de tiempo ARIMA.

## Introducción

Es conocido en el mundo, de los daños materiales y pérdidas humanas que ocasionan los sismos, por tanto, el problema de buscar predecir cuándo y dónde se va a producir constituye la preocupación de los diversos Institutos de investigación en muchos países en el mundo. Es por ello que diversas investigaciones apuntan contribuir en esta dirección. Hasta ahora se han ensayado diversas metodologías, diversos caminos, utilizando información de un tipo u otro, que van desde modelos físico matemáticos, modelos probabilísticos, hasta información de ondas electromagnéticas que se producen en

momentos previos a los sismos, etc. En el Perú, que forma parte de una de las zonas más sísmicas del Pacífico, esta preocupación señalada arriba también forma parte del IGP y de diversos investigadores. En el presente trabajo, que consideramos como un inicio en este campo, exploramos con diversas herramientas estadísticas patrones, en forma similar al que ya se aplicado en otras partes del mundo, en particular con series de tiempo. Previamente se han identificado las zonas sísmicas con mayor actividad.

En esta investigación, sobre el fenómeno de los sismos en el Perú, nos ha guiado la idea que podrían existir ciertos patrones de comportamiento aleatorio, en la ocurrencia de este fenómeno.

Sobre la búsqueda de estos patrones y dada su complejidad en el mundo existen diversas propuestas y teorías, las cuales vamos a presentar a continuación, extraído de El Committee on the Economic Benefits of improved seismic monitoring (7), que resume los principales métodos de pronóstico de sismos.

**“El método de brecha sísmica** es una aplicación de uso frecuente de este concepto de predicción sísmica a largo plazo y la identificación de potencial sísmico A lo largo de los límites de placas simples, como la falla de San Andreas, la mayor parte del movimiento de la placa, a largo plazo ocurre durante infrecuentes grandes terremotos”.

“Este enfoque fue exitoso y aplicado para varios terremotos grandes ( $M_w > 7.5$ ) a lo largo de las zonas de subducción y límites de placas de deslizamiento durante los años 60 y 70. Aplicación de este método para eventos con  $M_w < 7.5$ , o en áreas con configuraciones tectónicas complejas, es más controvertido”.

**“El patrón repetitivo de acumulación y liberación de tensión** también es expresado en patrones de sismicidad, donde los bajos niveles de sismicidad en la primera parte del ciclo (una vez que las réplicas del último evento disminuyan) son seguidos por un aumento en la actividad regional a medida que la placa reacumula y libera en última instancia, por la ocurrencia de otro terremoto con sus respectivos sismos previos y réplicas”.

**“Los métodos de reconocimiento de patrones,** se basan en estadísticas de cambios, tales como la tasa de ocurrencia de terremotos y la proporción de sismos grandes a pequeños, dentro de una región. Específicamente, el fenómeno que es monitoreado y analizado incluye pequeños sismos que se convierten más frecuentes, en un área que no es necesariamente donde el inminente terremoto ocurrirá; terremotos cada vez más agrupados en el tiempo y espacio; terremotos que ocurren casi simultáneamente a grandes distancias dentro de la región sísmica; y una proporción creciente de magnitud media a terremotos de pequeña magnitud. Estos métodos son utilizados para predecir la ocurrencia de un gran terremoto dentro de un periodo de meses a años, en una región circular grande especificada teniendo un radio de varios cientos de kilómetros”.

“El periodo de tiempo de sucesivas predicciones ha ido disminuyendo a medida que el desarrollo de los métodos han progresado. Los desarrolladores afirman haber predicho un número de terremotos que ocurrieron dentro de grandes regiones circulares: el 1989  $M_w 6.9$  Terremoto de Loma Prieta dentro de un periodo de 5 años, el



1994 Mw 6.7 terremoto de Northridge dentro de un periodo de 18 meses (aunque en realidad ocurrió 21 días después del final de la ventana), y el 2004 Mw 6.5 Terremoto de San Simón dentro de un periodo de 9 meses. Los desarrolladores predijeron que un terremoto de magnitud 6,4 o mayor ocurriría en el sudeste de California antes del 5 de septiembre de 2004, pero no ocurrió tal terremoto".(7)

La presente investigación se realiza, dado el interés público del tema sísmico, el mismo que puede servir para entender mejor el fenómeno y sus tendencias.

## Marco teórico y conceptual

Desde los años 60, China ha estado conduciendo sistemáticas observaciones, monitoreando anomalías de los sismos e intentando predecir los terremotos y la aplicación de evaluaciones, dependiendo del tiempo para disminuir los daños que éstos ocasionen. En tales esfuerzos, en algunos casos con éxitos y fracasos, sirven como lecciones en la ciencia de los terremotos, así como en la sostenibilidad social.

Para Pozos, Gómez y Hong (2014), el uso de redes neuronales artificiales es explorado para predecir aceleraciones máximas del terreno y pseudo aceleraciones para sismos de tipo intraslab e interplaca. Un total de 277 y 418 registros sísmicos de dos componentes para sismos de intraslab.

Hung y Nan (2003) utilizan el método bayesiano y usan un modelo epidémico para analizar datos de sismos en Taywan, siguiendo la idea de Ogata (1988). Han realizado estimaciones de riesgo de sismo por zonas con características especiales, y han realizado simulación mediante el método de Montecarlo. El ha usado el estimador de Máxima Verosimilitud (MLE), para analizar y predecir la actividad sísmica a lo largo de Nor-Este de la costa de Japón. La elección de la función de intensidad condicional ha sido descrito ampliamente en Ogata (1988) y varias referencias allí mencionadas.

También Peruggia y Santner (1996) (7) aplicó el método Bayesiano a datos sísmicos en el área de Sannio Matese en Italia, basado en la función de intensidad

$$\lambda(t | \underline{\theta}) = \mu + \sum_{t_i < t} e^{\beta(m_i - M_T)} \alpha e^{-\alpha(t-t_i)}$$

Donde  $\theta = (\mu, \alpha, \beta, a)$  son todos parámetros desconocidos.

Un método similar ha sido aplicado en el área de Hualien en el Nor-este de Taiwan por Fan y Li en el año 2002.

Alawadhi y Alhulail (2016) analizan datos de sismos en el El cinturón Zagros el cual es uno de los más sísmicamente activos del mundo. Debido a la ubicación de Kuwait en el suroeste del cinturón de Zagros, se ve afectada por movimientos tectónicos relativos en la región vecina. En su trabajo consideran fundamental evaluar la sismicidad de Zagros riesgos en Kuwait utilizando datos recientes y coordinar con los organismos respectivos para reducir dichos riesgos. Utilizando la magnitud de la onda del cuerpo (Mb) recogidos en Kuwait, evalúan los cambios recien-

tes en la magnitud de los terremotos y sus variaciones en la vecindad de Kuwait. Han construido un modelo de punto de cambio para detectar los cambios significativos en sus parámetros.

Volpe y Piersanti (2016) analizan la posible relación entre los dos sismos ocurridos en Emilia Romagna, Italia, el 20 y 29 de mayo de 2012. Realizan simulaciones en 3D, con datos satelitales de las ondas de los sismos. Asimismo, investigan las posibles conexiones entre las principales Acontecimientos de la secuencia sísmica Emilia del 2012 que ocurrieron los días 20 y 29 de mayo, con el fin de aclarar el papel que la Transferencia de estrés y migración fluida profunda en la ocurrencia del terremoto del 29 de mayo después del primer sismo.

## Algunos conceptos

**Sismo**, también conocido como terremoto, se denomina una sacudida de la tierra que consiste en una serie de vibraciones en la superficie como consecuencia del movimiento de las capas interiores de la tierra.

Localidad más cercana al sismo. Es la localidad que según el IGP esta cerca al epicentro del sismo.

**La escala sismológica de Richter** (13) o escala de magnitud local (ML), es una escala logarítmica arbitraria que asigna un número para cuantificar la energía liberada en un terremoto, denominada así en honor del sismólogo estadounidense Charles Richter (1900-1985). Escala de Richter.

**Fecha del sismo**, se refiere a la fecha de la ocurrencia del sismo

**La latitud** es la distancia angular entre la línea ecuatorial (el ecuador), y un punto determinado de la Tierra, medida a lo largo del meridiano en el que se encuentra dicho punto. Según el hemisferio en el que se sitúe el punto, puede ser latitud norte o sur.

**La altitud**

**Magnitud del sismo**, La magnitud de un sismo es un número que busca caracterizar el tamaño de un sismo y la energía sísmica liberada. Se mide en una escala logarítmica, de tal forma que cada unidad de magnitud corresponde a un incremento de raíz cuadrada de 1000, o bien, de aproximadamente 32 veces la energía liberada

**Intensidad local**, es la intensidad del sismo en referencia a la localidad o localidades mas cercanas.

## Metodología

El presente estudio es de tipo exploratorio en la medida en que el fenómeno sísmico es muy complejo y aleatorio y a través de los resultados se pretende conocer sus características y tendencias que serán expresadas en el modelo o los modelos estadísticos.

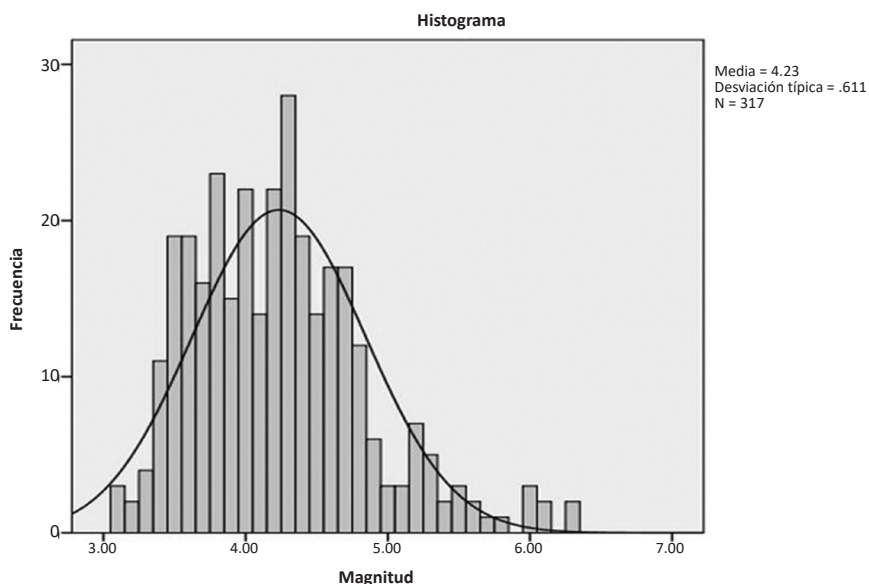
Por otro lado se utilizó información secundaria, los que proporciona el Instituto Geofísico del Perú y otras fuentes, disponibles.

Para realizar el análisis se emplearon diversa herramientas estadísticas, tales como el Cluster Analysis, modelos de Series de Tiempo ARIMA.

Se utilizó el software estadístico SPSS y R.

## Resultados y aportaciones

Grafico 1  
Distribución de las magnitudes de los sismos 2017



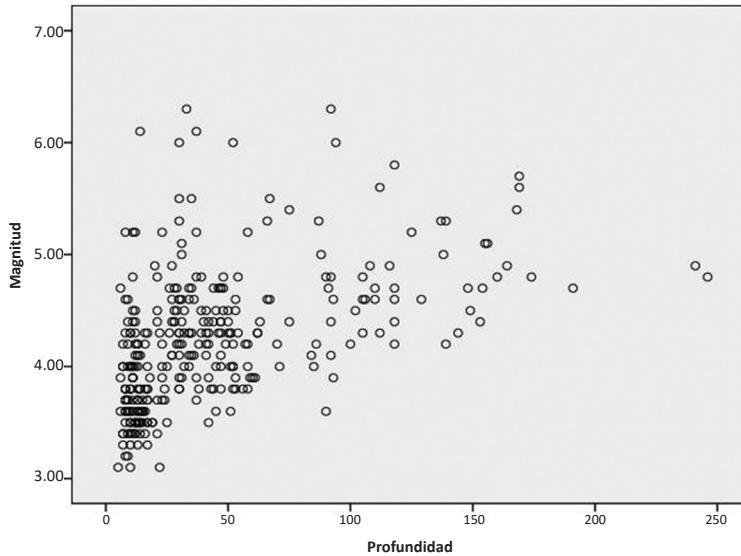
Fuente: Elaboración propia (Datos del IGP magnitud de los sismos de Enero a Noviembre 2017).

Se puede observar que los sismos en el Perú, se distribuyen aproximadamente como una normal, según el gráfico y el cuadro 1-A del Anexo, que muestra la prueba de Kolmogorov – Smirnov para la distribución normal.

## Relación entre la magnitud y la profundidad del sismo

Gráfico 2

Relación entre la magnitud y la profundidad



Fuente: Elaboración propia

Tabla 1

		Profundidad
magnitud	Correlación de Pearson	0.492
	Sig. (bilateral)	0
	N	317

Fuente: Elaboración propia

La correlación entre la magnitud de los sismos y la profundidad no es alta, como se muestra en la tabla 1. Tampoco esta correlación lineal no es tan buena (tabla 2).

Tabla 2

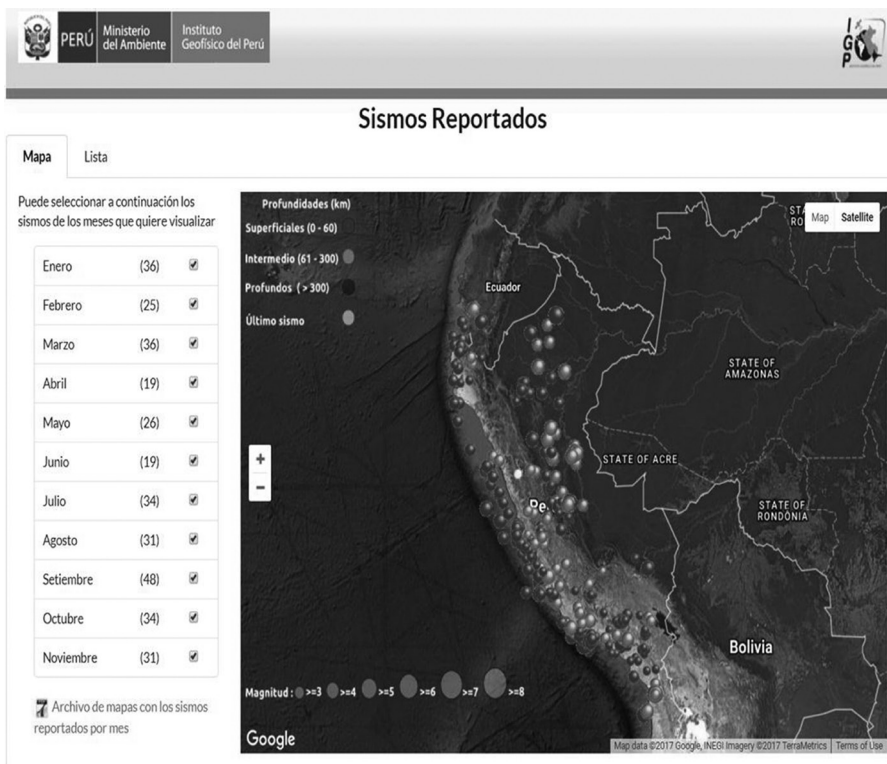
Descriptivos para la serie magnitud						
	N	Rango	Mínimo	Máximo	Media	Desv. típ.
magnitud	317	3.20	3.10	6.30	4.2341	.61137
Profundidad	317	241	5	246	44.13	42.954
N	317					

Fuente: Elaboración propia

A nivel nacional el promedio de la magnitud de los sismos de enero a noviembre del 2017 fue de 4.23 en la escala de Richter y la profundidad fue de 44.13 Km.

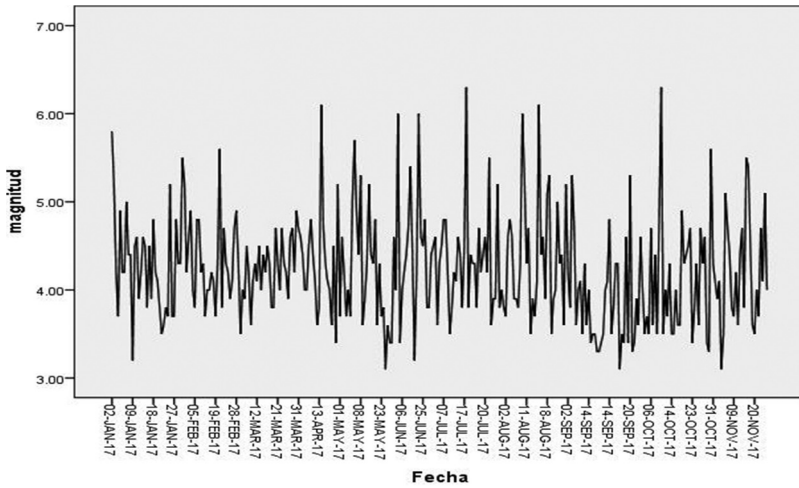
Gráfico 3

**Reporte de los sismos del 2017 por parte del Instituto Geofísico del Perú (IGP)  
30-11-17**



En este mapa de los sismos reportados por el IGP, se puede observar la mayor concentración en la zona de Arequipa, seguido de Lima-Ica.

Gráfico 4  
Serie de la magnitud de los sismos del 2017



Fuente: Elaboración propia.

En este gráfico se puede observar la oscilación de la magnitud de los sismos, de enero a noviembre de 2017.

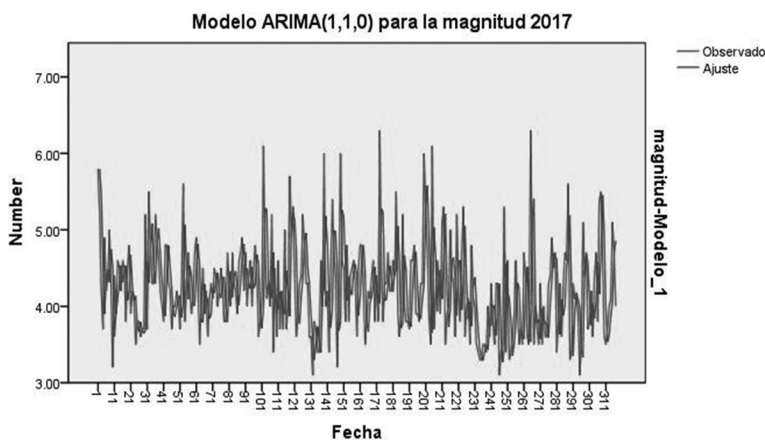
Luego de ensayar con varios modelos de series de tiempo, hemos encontrado el modelo ARIMA(1,1,0), el cual se ajusta mejor a los datos de la magnitud de los sismos en el Perú.

Tabla 3

Estadísticos del modelo ARIMA (1,1,0)					
Modelo	Estadísticos de ajuste del modelo		Ljung-BoxQ(18)		Número de valores atípicos
	R-cuadrado estacionaria	Estadísticos	GL	Sig	
magnitud	.166	54.076	17	.000	0

Fuente: Elaboración propia

Gráfico 5  
La magnitud observada y la ajustada



Fuente: Elaboración propia

Este gráfico se puede observar un cierto buen ajuste entre la serie observada y la serie ajustada. Lo anterior nos plantea la siguiente interrogante, existe un patrón de comportamiento de la serie de los sismos en el Perú en el 2017?

### Análisis de grupos y sus características

Para hallar los conglomerados, se ha utilizado la herramienta Cluster Analisis, con distancia euclidea entre los grupos, y con la ayuda del grafico Dendograma, detectamos los 8 grupos.

Tabla 4  
Características de los sismos por grupo (2017)

Grupo	Nº de sismos	Ciudad	Magnit prom	Prof prom	Desv. Tip Mag	Desv. Tip Prof
1	14	Moyobamba	4.81	72.79	0.66432	63.011
2	7	Tacna	4.8	62.14	0.92556	38.321
3	91	Lima - Ica	4.07	41.87	0.45088	28.933
4	127	Arequipa	4.04	32.9	0.59831	41.694
5	19	Chimbote	4.59	45.05	0.34302	15.629
6	32	Pucallpa	4.66	80.25	0.57283	67.431
7	25	Sullana	4.47	43.8	0.52144	19.17
8	2	San Gaban	4.5	12	0.98995	5.657
Total	317					

Fuente: Elaboración propia.

Respecto a la magnitud de los sismos, Moyobamba y Tacna presentan los mayores promedios, sin embargo se puede observar que no difieren de los demás grandemente. También son los mismos quienes presentan mayor varianza.

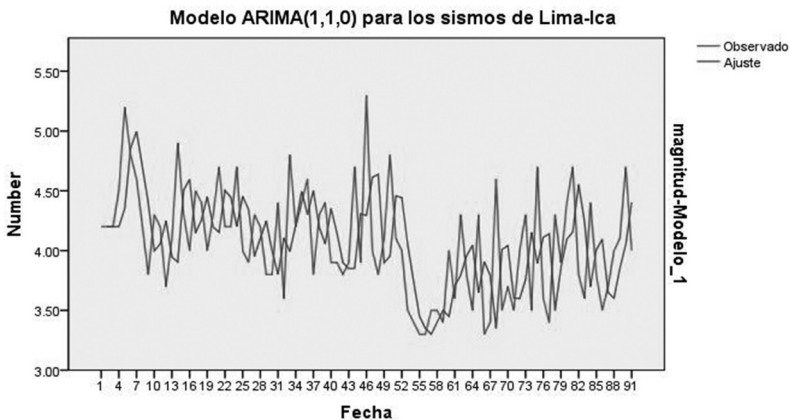
### 1. Un análisis para el grupo de Lima - Ica

Tabla 5

Estadísticos del modelo ARIMA (1,1,0)					
Modelo	Estadísticos de ajuste del modelo		Ljung-BoxQ(18)		Número de valores atípicos
	R-cuadrado estacionaria	Estadísticos	GL	Sig	
magnitud	.242	34.557	17	.007	0

Fuente: Elaboración propia

Gráfico 6



Fuente: Elaboración propia

En el caso de los sismos del grupo de Lima-Ica, se observa que a partir de la mitad de año, (septiembre) la tendencia es a una disminución de la media de la serie. El ajuste es relativamente bueno, sin embargo que se ajuste a un modelo requiere de un análisis más especializado desde el punto de vista del sismólogo?



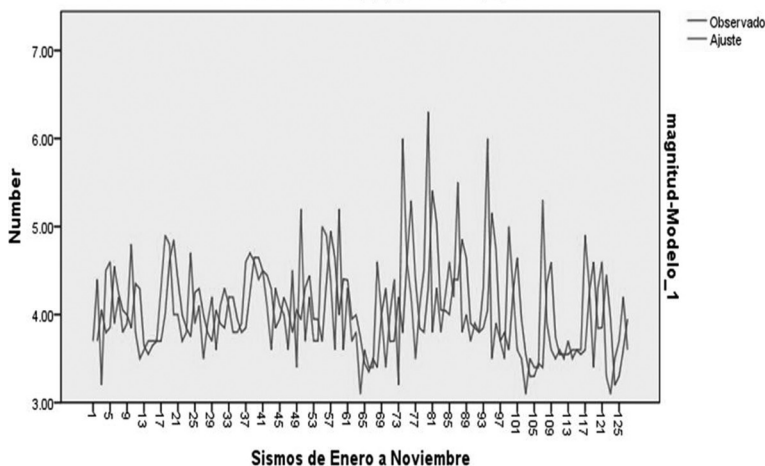
## 2. Un análisis para el grupo de Arequipa

Tabla 6

Estadísticos del modelo ARIMA (1,1,0)					
Modelo	Estadísticos de ajuste del modelo		Ljung-BoxQ(18)		Número de valores atípicos
	R-cuadrado estacionaria		Estadísticos	GL Sig	
magnitud	.429		39.471	17 .002	0

Fuente: Elaboración propia

Gráfico 7  
Modelo ARIMA(1,1,0) Para Arequipa



Fuente: Elaboración propia

En el caso de Arequipa, no se presenta la tendencia a disminuir la media como es el caso de Lima-Ica, pero es parecido en el sentido del relativo buen ajuste con el modelo ARIMA(1,1,0).

### Discusión de resultados

Dada la complejidad de los datos de los sismos en el Peru, sin embargo nos ha guiado la búsqueda de patrones de comportamiento, de los resultado obtenidos en la sección anterior se puede observar lo siguiente:

- Para el conjunto de datos de los sismos del 2017, se ha constatado una relación lineal significativa entre la profundidad del sismo y la magnitud.

- Asimismo, se ha ajustado a los datos de la magnitud de los sismos, un modelo ARIMA(1,1,0), el cual ha resultado significativo, tanto a nivel nacional, así como para el caso de Lima-Ica y Arequipa.
- También se ha encontrado ocho grupos de zonas sísmicas, evaluadas por las distancias. Siendo el grupo de Arequipa, seguido de Lima-Ica, como los que presentan mayor actividad sísmica.
- Para el conjunto de datos de los sismos del grupo de Lima-Ica
- Se ha encontrado una tendencia lineal negativa de la magnitud de los sismos.
- Y respecto a la profundidad, son Pucallpa y Moyobamba, quienes presentan la mayor profundidad promedio 80,25 km y 72,29 km, respectivamente, y también la mayor varianza, 67 km y 63 km.
- Es la selva la que presenta la mayor profundidad y la mayor magnitud promedio.
- Hemos encontrado que la magnitud promedio de Arequipa y de Lima-Ica son bastantes parecidos 4,04 y 4,07 y su profundidad promedio no difieren grandemente, 32,9 km y 41,87 km, respectivamente.

## Conclusiones

1. Hemos encontrado ocho zonas de actividad sísmica diferenciados, de las cuales Arequipa y Lima son las más activas.
2. El modelo de series de tiempo ARIMA(1,1,0) de las magnitudes de los sismos de Enero a Noviembre del 2017, ha resultado significativo a nivel nacional, y también ha sido significativo para Lima-Ica y Arequipa.
3. En el grupo de Lima, existe una tendencia hacia una pequeña disminución en el nivel de la media de la serie, lo que podría estar indicando la cercanía de un sismo de mayor magnitud en los próximos meses.

## Referencias bibliográficas

- Tsay Hung Fan and Eng Nan Kuo " A Bayesian analysis for seismic data on Taywan. 2003.
- Pozos-Estrada, A, Gómez, R., y Hong, H.P (2014). "Use of Neural network to predict the peak ground accelerations and pseudo spectral accelerations for Mexican Inslab and Interplate Earthquakes".
- Fuqiong Huang\*, Mei Li, Yuchuan Ma, Yanyan Han, Lei Tian, Wei Yan, Xiaofan Li. "Studies on Earthquake precursors in China", 2016.
- Pecker, A., Fassioli, E. An overview on the research SIGMA Proyect. Edit Springer 2017.

Ogata 2005, Anomalies of Seismic Activity and Transient Crustal Deformations Preceding the 2005 M 7.0 Earthquake West of Fukuoka.

Committee on the Economic Benefits of Improved Seismic Monitoring Improved seismic monitoring decision making. Assessing the Value of Reduced Uncertainty The National Academies press. Washington, D.C. 2006

“Escala de Richter”. En: Significados.com. Disponible en: <https://www.significados.com/escala-de-richter/>

Douglas C. Montgomery , Cheryl L. Jennings , Murat Kulanci, 2015, Introduction To Time Series Analysis and Forecasting.

John K. Kruschke, Doing Bayesian Data Analysis: A Tutorial with R and BUGS, 2010

Mike West, Jeff Harrison, Bayesian Forecasting and Dinamic Models. 1989.

Volpe, M. y Piersanti, A. (2016)The role of fluid migration and static stress transfer in searching connections between the May 2012 Emilia earthquakes through a fully 3D finite element modeling. DOI 10.1007/s40808-016-0087-9

## ANEXO 1

Este tipo de gráficos, nos sirve para identificar que tipo de modelo de series de tiempo podemos ensayar ajustar a la serie de la magnitud de los sismos.

Tabla 1-A

Prueba de Kolmogorov-Smirnov para una muestra		
		magnitud
N		317
Parámetros	Media	4.2341
normales <sup>a,b</sup>	Desviación típica	0.61137
	Absoluta	0.082
Diferencias más extremas	Positiva. Positiva más extremas	0.082
	Negativa	-0.058
Zde Kolmogorov-Smirnov		1.454
Sig. asintót. (bilateral)		0.029

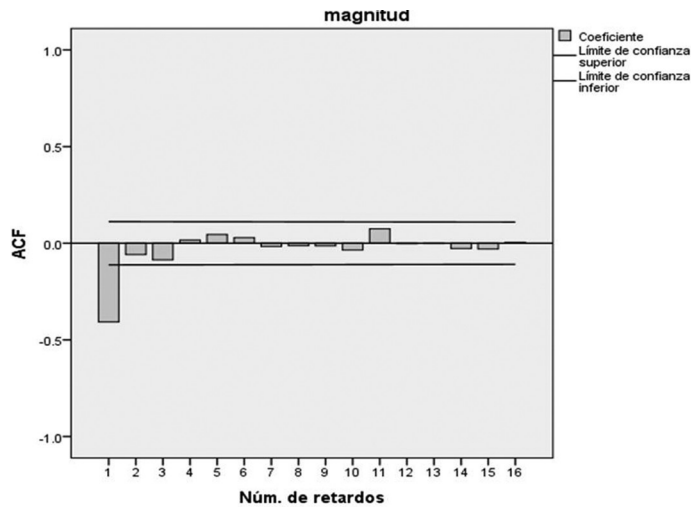
a. La distribución de contraste es la Normal.  
b. Se han calculado a partir de los datos.

Fuente: Elaboración propia

## Un análisis de series de tiempo para la magnitud a nivel nacional

Gráfico A-1

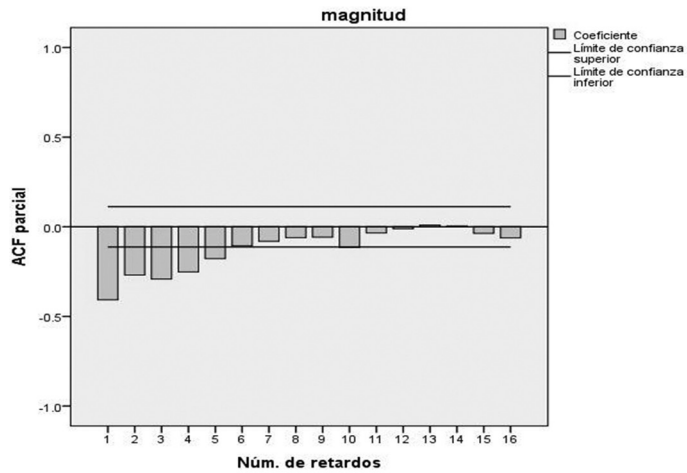
### La función de autocorrelación de la magnitud



Fuente: Elaboración propia

Gráfico A-2

### La función de autocorrelación parcial de la magnitud a nivel nacional



Fuente: Elaboración propia

Tabla A- 2  
**Descriptivos por grupos**

<b>Grupos</b>	<b>Variable</b>	<b>N</b>	<b>Mínimo</b>	<b>Máximo</b>	<b>Media</b>	<b>Desv. típ.</b>
Moyobamba	magnitud	14	3.60	6.10	4.8143	.66432
	Profundidad	14	13	174	72.79	63.011
Tacna	magnitud	7	3.60	6.30	4.8000	.92556
	Profundidad	7	9	110	62.14	38.321
Lima-Ica	magnitud	91	3.30	5.30	4.0736	.45088
	Profundidad	91	6	129	41.87	28.933
Arequipa	magnitud	127	3.10	6.30	4.0402	.59831
	Profundidad	127	5	246	32.92	41.694
Chimbote	magnitud	19	4.10	5.50	4.5895	.34302
	Profundidad	19	27	91	45.05	15.629
Pucallpa	magnitud	32	3.90	6.10	4.6656	.57283
	Profundidad	32	8	191	80.25	67.431
Sullana	magnitud	25	3.80	6.00	4.4760	.52144
	Profundidad	25	11	93	43.80	19.170
San Gaban	magnitud	2	3.80	5.20	4.5000	.98995
	Profundidad	2	8	16	12.00	5.657

*Fuente: Elaboración propia.*

# Competencias laborales y demanda laboral de los egresados de Ingeniería Económica, UNI 2017

José Saavedra Pacheco

## Resumen

El presente estudio sobre la relación entre las competencias laborales de los egresados de la Escuela Profesional de Ingeniería Económica. y la demanda de las empresas, analiza cuantitativamente el proceso de inserción laboral de los egresados universitarios de dicha escuela conectado a los requerimientos que las empresas privadas o públicas en Lima metropolitana que solicitan en la práctica del ejercicio profesional del economista. Este estudio, pone de manifiesto la sensibilidad e importancia que la EPIEC\_ UNI, debe poner en la formación para el empleo de sus titulados. El trabajo de investigación se divide en 4 capítulos:

El capítulo 1 se refiere al planteamiento del problema en dos aspectos: La fundamentación del problema, en el cual se determina que existe un divorcio entre lo que la escuela profesional de ingeniería económica provee y lo que exigen los empleadores. En segundo lugar se plantea la pregunta de investigación: ¿Cuáles son las competencias laborales que los egresados de la escuela profesional de ingeniería económica. UNI deben adquirir para insertarse con éxito en el mercado laboral de Lima metropolitana?

En el capítulo 2, marco teórico, se destaca el concepto de competencias laborales entendidas como las capacidades que requieren los egresados para afrontar con éxito las diferentes áreas del quehacer de las empresas, en el campo macroeconómico, microeconómico, marketing y proyectos de inversión,

El capítulo 3, presenta las hipótesis del estudio. La hipótesis principal: Si los egresados de ingeniería económica de la UNI adquieren competencias laborales que el mercado laboral exige, se insertarían con mayor éxito. Se presentan las hipótesis secundarias

El capítulo 4, presenta la metodología y el desarrollo de la investigación: La opinión de los egresados de los cuatro últimos ciclos académicos mediante encuestas y por otro lado la selección de 20 empresas, privadas y del sector público, para recabar la versión del empleador respecto a las capacidades de los egresados de la EPIEC- UNI

El capítulo 5 presenta las principales conclusiones y encuentra una relación significativa entre las competencias laborales de los egresados de la EPIEC-UNI y la demanda laboral.

**Palabras clave:** Competencias laborales, demanda laboral, inserción laboral de egresados, competencias específicas.

## Introducción

La educación superior siempre ha sido una parte esencial de las sociedades, dado que produce individuos altamente educados que poseen los conocimientos para realizar desarrollo e innovación y que pueden llevar el país a un nivel superior en el campo de investigación, liderazgo y producción económica. Por décadas los gobiernos e instituciones de educación superior dieron la empleabilidad de graduados de educación superior por supuesto, porque formaban un grupo privilegiado que hizo fluidamente la transición al mercado laboral y rindió buenos resultados en términos de remuneración y desarrollo profesional. Aunque ya en los años 70 había empleadores que quejaron que los graduados no poseían las capacidades que necesitaban, esto no resultó en situaciones de desempleo o sobre-educación entre graduados, porque se necesitaron a todos los egresados universitarios para cumplir con la demanda (Mason et. al., 2003: 3). Sin embargo, en las décadas después ocurrió una expansión impresionante de la educación terciaria, por lo que la oferta de personal altamente educado aumentó rápido y los graduados empezaron a sentir la presión de la competencia. Los empleadores, que ahora disponían de suficiente capital humano, pudieron hacer una selección de los graduados en base de sus competencias, por lo que parte de los titulados se quedaban al margen del mercado laboral. Además de la expansión de la educación superior, ocurrieron algunos otros desarrollos que influyeron la posición de los graduados, como el surgimiento de nuevas tecnologías (sobre todo en el campo de la informática) y formas de trabajo, la globalización, y la creciente movilidad de personal y estudiantes.

Hoy en día existe un debate mundial sobre la empleabilidad de graduados y, en general, sobre el objetivo de la educación superior con relación al mercado laboral. Dado que la transición al mercado laboral se ha puesto más compleja y duradera por graduados, se discute la necesidad de modificar los programas de educación superior y adaptarlos mejor a los requisitos de los empleadores.

En la presente investigación se analiza la posición actual de los graduados de ingeniería económica uni y su relación con el mercado laboral. El segundo capítulo investiga la inserción de los graduados en el mercado laboral y la dinámica del juego de oferta

y demanda de personal académico. Después, en el tercer capítulo se analiza la calidad de la posición de los graduados en el mercado laboral evaluada por medio de las competencias que requiere los empleadores en Lima Metropolitana. Finalmente, en el cuarto capítulo se hace una comparación entre las competencias de los graduados y las expectativas de los empleadores, en base de que se hacen conclusiones sobre la empleabilidad real de los graduados y el valor profesional de una formación universitaria.

## Contexto teórico

A fines del siglo XX, el concepto de competencia se había extendido, tanto en los países desarrollados como en América Latina, especialmente en los ámbitos laboral y educativo. Sin embargo, era evidente que existían dos grandes tendencias para aproximarse a dicho concepto: la del mercado de trabajo, que está asociada a las denominadas competencias profesionales o laborales, y la educativa, vinculada a las competencias de egreso o curriculares. Antes de explicar dicha división, conviene revisar los antecedentes que comparten ambas tendencias. En ese sentido, es posible identificar diversas propuestas; de estas, en un primer lugar, conviene presentar a las que aluden a la etimología del término. Así, Weigel *et al.* (2007) ubican las raíces del concepto en la época de la Grecia Clásica y señalan que el equivalente de la palabra competencia en el antiguo griego es “*ikanótis*”. La misma que se traduce como la cualidad de ser “*ikanos*” (capaz), tener la habilidad de conseguir algo, destreza. Asimismo, mencionan que para nombrar capacidad o competencia profesional/vocacional, se usa el término griego “*epangelmatikes ikanotita*”.

Por su parte, Peñaloza (2000) indica que las palabras “competente” y “competencia” provienen, respectivamente, del latín “*competens*”, concebido como el ser capaz, y “*competentia*”, entendido como la capacidad y la permisión. A su vez, estos vocablos se habrían formado de los términos “*cum*” y “*peto*”. El primero significa “con”; mientras que el segundo correspondería al infinitivo de “*petere*”, cuyo significado es “ir” y “atacar”. De este modo, “*cumpetere*” quiere decir, por un lado, “ir con”, “ir a la par con otro” o “ser adecuado a algo”; pero también puede entenderse como “luchar con” o “atacar a otro”. Asimismo, Levy-Leboyer (1998) plantea que la raíz latina también da lugar a dos acepciones en el idioma español: “*competer*” y “*competir*”. En ese sentido, mientras que el sustantivo *competencia* se relaciona con los dos verbos; en cambio, el adjetivo “*competente*” solo se relaciona con el primero. Para el siglo XVI, dicho término era reconocido en Europa Occidental; específicamente, en los idiomas inglés, francés y holandés; con el uso de las palabras “*competence*” y “*competency*” (Weigel *et al.* 2007).

Por otro lado, de las dos tendencias que se extendieron durante el siglo XX, la primera es la relacionada con el mercado de trabajo. Respecto del origen de la misma existen diversas propuestas. Así, para Tobón (2006a), el concepto de competencia se comenzó a estructurar en la década de 1970 sobre la base de dos aportaciones: la



lingüística de Noam Chomsky y la psicología conductual de Skinner. El primer autor propuso el concepto de “competencia lingüística” como una estructura mental implícita y genéticamente determinada que se ponía en acción mediante el desempeño comunicativo; se trataba pues del uso efectivo de la capacidad lingüística en situaciones específicas (Chomsky, 1970). Asimismo, señaló que la competencia lingüística es un dispositivo natural que permite el aprendizaje de la lengua materna de acuerdo con los contextos donde interactúan los sujetos. Para ello, utilizó dos términos: “*performance*” y “*competence*”. El primero se refiere a la comunicación y creación del lenguaje; mientras que el segundo alude al dispositivo para la adquisición de la gramática de una lengua (Chomsky, 1972). Dicho concepto comenzó a tener múltiples desarrollos, críticas y reelaboraciones, tanto en la lingüística como en la psicología (conductual y cognitiva) y en la educación (Tobón, 2006a). Contraria a la propuesta de Chomsky, quien considera a la competencia como un fenómeno interno; está la línea conductual de Skinner, para quien la competencia es un comportamiento efectivo. En ese sentido, se ha desarrollado y extendido un sólido modelo conductual de las competencias que se basa en el comportamiento observable, efectivo y verificable; lo cual ha generado que el desempeño confluya dentro de la competencia. Así, el enfoque de competencias, desde la línea conductual, ha alcanzado notables desarrollos en el campo de la gestión del talento humano en las Organizaciones, donde se busca que los trabajadores posean competencias clave para que las empresas sean competitivas (Tobón, 2006a).

Por su parte, la psicología cultural enfoca a la competencia como un concepto base de la interacción de la persona con el entorno. Según Torrado (2000), Vygotsky (1896-1934) plantea que el desarrollo cognitivo se explica por el impacto del mundo social y cultural del sujeto dentro de su actividad psicológica. Para la autora, ser competente se traduce en la acción que surge en un contexto particular, más allá de los conocimientos formales. Es decir, competencia es la capacidad de realización, situada y afectada por y en el contexto donde se desenvuelve el sujeto. Así, la psicología cultural ha aportado al concepto de competencias el principio de que la mente y el aprendizaje son una construcción social y requieren de la interacción con otras personas, que se ven influenciadas por el contexto (Tobón, 2006a). Asimismo, desde el punto de vista de la psicología cognitiva, se destaca el desarrollo de la Teoría de Inteligencias Múltiples propuesto por Gardner como aporte a la definición de competencias. El autor señala que “*la inteligencia es un conjunto de rasgos que se pueden evaluar confiablemente por medio de las competencias intelectuales, los procesamientos de la información, las capacidades y habilidades cognitivas y las formas de conocimiento*” (Gardner, 1994: 394). Dichas competencias intelectuales pueden apreciarse tanto en el desarrollo de actividades cotidianas como actividades más innovadoras relacionadas con el arte, la ciencia, el trabajo, etc. Es decir, que las competencias intelectuales o cognitivas evolucionan a medida de que el sujeto se interrelacione en un contexto sociocultural, a través de símbolos que le permitirán reconocer, interpretar y conectarse con

su espacio (Rodríguez, 2007). Según Guerrero (1999), el enfoque de competencias profesionales tiene doble rige: normativo y empresarial. Por un lado, está unido a la regulación de las cualificaciones profesionales de los trabajadores en ejercicio para su posible normalización. Y de otro, aparece ligado a las políticas de gestión de personal que llevan a cabo consultorías y empresas *ad hoc* (grandes compañías), con el objetivo permanente de rentabilizar, optimizar o adecuar la mano de obra a sus necesidades cambiantes, atribuidas a los cambios en la organización del trabajo y del consumo. Según el pedagogo alemán Gerhard Bunk (1994), el concepto procede del campo de la organización y se refiere tanto a la regulación de las atribuciones en el seno de las organizaciones, como a la facultad de toma de decisiones que tienen los diferentes titulares de las empresas. El enfoque orientado a la competencia es un “cambio paradigmático”, un “paso cualitativo” que supera, en los objetivos de la formación profesional, al concepto de cualificación. La evolución de tales objetivos: concepto de “capacidades profesionales” situado con anterioridad a los 60; el de “cualificaciones profesionales”, entre los 60 y 70; y el de la “competencia profesional”, a la que considera “como objetivo global del proceso de aprendizaje”, a partir de los 80. A pesar de los importantes aportes y propuestas anteriores, se tiende a reconocer que el origen de las competencias, desde la perspectiva del mercado de trabajo, se encuentran en los trabajos de McClelland de la Universidad de Harvard, quien, durante la década de 1970, realizó investigaciones en el área motivacional y laboral. En función de los hallazgos de sus investigaciones, planteó que las pruebas de desempeño tradicionales no eran efectivas para predecir el éxito en el trabajo. Para explicarlo, propuso que se debían considerar las características de los trabajadores y sus comportamientos concretos ante situaciones de trabajo, en lugar de los expedientes académicos y los coeficientes de inteligencia. El autor propuso seis factores que son predictores universales del éxito profesional y que son independientes del género, raza o situación social: espíritu de logro, de servicio, capacidad de influir, de gestionar, de solucionar problemas y eficacia personal. A estos factores los llamó competencias, incluyendo por primera vez el concepto dentro de la gestión de las personas (McClelland, 1973). Así, a partir de 1980, el concepto se desarrolla básicamente en el ámbito laboral, al ritmo de los diversos cambios que ocurren a nivel mundial en torno al desarrollo científico y tecnológico. En consecuencia se obliga a las naciones y organizaciones a repensar los procesos relacionados con la producción, la estructura y funciones del trabajo, así como también la formación indispensable para ser competitivos en un mundo cada vez más globalizado (Rodríguez, 2007).

## Las competencias laborales

Existen múltiples y variadas definiciones en torno a la competencia laboral. Un concepto generalmente aceptado la establece como una capacidad efectiva para llevar a cabo exitosamente una actividad laboral plenamente identificada. La competencia laboral no es una probabilidad de éxito en la ejecución del trabajo, es una capacidad

real y demostrada. A continuación, podrá consultar varias definiciones sobre competencia laboral:

Enfoque conductista: CONOCER (México)(1): capacidad productiva de un individuo que se define y mide en términos de desempeño en un determinado contexto laboral, y no solamente de conocimientos, habilidades, destrezas y actitudes; estas son necesarias pero no suficientes por sí mismas para un desempeño efectivo.

INEM (España) (2): “las competencias profesionales definen el ejercicio eficaz de las capacidades que permiten el desempeño de una ocupación, respecto a los niveles requeridos en el empleo. “Es algo más que el conocimiento técnico que hace referencia al saber y al saber-hacer”. El concepto de competencia engloba no solo las capacidades requeridas para el ejercicio de una actividad profesional, sino también un conjunto de comportamientos, facultad de análisis, toma de decisiones, transmisión de información, etc., considerados necesarios para el pleno desempeño de la ocupación. POLFORM/OIT (3): La competencia laboral es la construcción social de aprendizajes significativos y útiles para el desempeño productivo en una situación real de trabajo que se obtiene no solo a través de la instrucción, sino también –y en gran medida– mediante el aprendizaje por experiencia en situaciones concretas de trabajo. Cabe mencionar que la OIT ha definido el concepto de “Competencia Profesional” como la idoneidad para realizar una tarea o desempeñar un puesto de trabajo eficazmente por poseer las calificaciones requeridas para ello (4). En este caso, los conceptos competencia y calificación, se asocian fuertemente dado que la calificación se considera una capacidad adquirida para realizar un trabajo o desempeñar un puesto de trabajo. La comunidad Latina de estudiantes de negocios Los documentos que buscas están en <http://www.gestiopolis.com/> Provincia de Quebec: Una competencia es el conjunto de comportamientos socioafectivos y habilidades cognoscitivas, psicológicas, sensoriales y motoras que permiten llevar a cabo adecuadamente un papel, una función, una actividad o una tarea. Consejo Federal de Cultura y Educación(Argentina): Un conjunto identificable y evaluable de conocimientos, actitudes, valores y habilidades relacionados entre sí que permiten desempeños satisfactorios en situaciones reales de trabajo, según estándares utilizados en el área ocupacional AUSTRALIA (5): La competencia se concibe como una compleja estructura de atributos necesarios para el desempeño de situaciones específicas.

## Concepto de mercado laboral

Se denomina mercado de trabajo o mercado laboral al mercado en donde confluyen la demanda y la oferta de trabajo. El mercado de trabajo tiene particularidades que lo diferencian de otro tipo de mercados (financiero, inmobiliario, de materias primas, etc.) ya que se relaciona con la libertad de los trabajadores y la necesidad de garantizar la misma. En ese sentido, el mercado de trabajo suele estar influido y regulado por el Estado a través del derecho laboral y por una modalidad especial de contratos, los convenios colectivos de trabajo.

En el mercado de trabajo los individuos intercambian servicios de trabajo. La oferta de trabajo incluye al conjunto de trabajadores que están dispuestos a trabajar y la demanda está formada por el conjunto de empresas o empleadores que contratan a los trabajadores. El salario es el precio del trabajo realizado por el trabajador en este mercado.

La información sobre las tendencias del mercado laboral actual y de los sectores que generan empleo es básica para situarnos en la realidad del mundo del trabajo.

Suele estar influido y regulado por el Gobierno a través del derecho laboral y por una modalidad especial de contratos, los convenios colectivos de empleo.

Es importante conocer el funcionamiento del mercado laboral de la zona y del comportamiento de las ocupaciones para poder tomar una decisión en relación al objetivo profesional que se busque.

El mercado laboral ha sufrido cambios importantes en los últimos años, tanto desde el aspecto sociológico, como desde el económico y el tecnológico. Estos cambios han propiciado la desaparición de algunos trabajos y la aparición de otros. Los denominados yacimientos de empleo son los ámbitos dónde se encuentran las ocupaciones en alza.

Conocer la situación actual del mercado laboral y las previsiones de su evolución beneficia en el momento de fijar objetivos profesionales.

Toda organización en la medida en que ofrece oportunidad de trabajo, es parte integrante del mercado laboral, la convivencia entre el mercado laboral y la legislación laboral no siempre es posible pues, existen fuerzas que desean aumentar, restricciones impuestas por la legislación laboral al mercado laboral para que la sociedad sea más solidaria.

El mercado laboral es dinámico y experimenta continuos cambios, las características estructurales del mercado laboral influyen en las prácticas de RH de la empresa. Cuando predomina la demanda en el mercado laboral (las oportunidades de trabajos son mejores que la demanda).

El crecimiento de las empresas, aún en los escenarios más complicados, no se detiene. Probablemente se reducirá la velocidad del desarrollo, pero el imperativo es continuar. Entonces el mercado laboral tendrá su impacto, abriéndose nuevas plazas de empleo para la población económicamente activa.

La tasa del desempleo es un problema lacerante en la realidad ecuatoriana. Las Oportunidades son escasas y no todos pueden ocuparse.

Los requerimientos de las empresas también se vuelven cada vez más exigentes. No es suficiente contar con acreditaciones académicas, sino además mostrar buenas actitudes personales. De acuerdo a la información proporcionada por la división de selección de ejecutivos de Deloitte, durante el año 2009 y lo que va del presente, las áreas donde se demandan más personal son:

Operaciones

Mercadeo (publicidad, trade, merchandising, comunicación)

Ventas (cuentas claves, retail)

Finanzas (costos, presupuestos, controlling, análisis financiero, tesorería)

Auditoría

Recursos humanos

(Tomado de <http://mercadolaboraledu.blogspot.pe/2011/07/se-denomina-mercado-de-trabajo-o.html>)

## **Mercado laboral, oferta y demanda laboral**

“El mercado laboral es el entorno en el que se interrelacionan las ofertas de trabajo y las demandas de empleo. Por tanto, constituye el ámbito en el que se define y desarrolla todo lo relacionado con las ocupaciones: empresas, personas que buscan empleo, contratos instituciones, etc.” (Santomé, 2004; 1)

El mercado laboral se define como aquel mercado específico que se compone de dos variables fundamentales: la oferta del mercado laboral y la demanda del mercado laboral. (Fontana, 2003; 122)

Ambos autores coinciden en que el Mercado Laboral es aquel espacio en el se interrelacionan dos variables fundamentales, las cuales son la oferta y la demanda. Concluimos entonces que el mercado de trabajo busca asignar los recursos humanos interesados en trabajar (oferta laboral) en las plazas laborales generadas por las empresas (demanda de mano de obra).

### **Oferta del mercado laboral**

La oferta laboral está constituida por la oferta de mano de obra que es reclamada por las empresas en función de sus necesidades productivas. Dicho de otro modo, la oferta laboral se constituye por el conjunto de trabajadores que ofrecen su fuerza de trabajo a cambio de un salario. (Fontana, 2003; 122)

La oferta del mercado de trabajo o también llamada oferta agregada, está conformada por el conjunto de las ofertas de trabajo existentes en un momento dado. (Calvo, 2005; 39)

Podemos notar claramente que ambas definiciones son contrarias, la primera hace referencia a que la oferta laboral se constituye por el número de personas que desean trabajar a cambio de un salario; la segunda por el contrario afirma que las ofertas laborales se conforman por los puestos de trabajo que ofrecen las empresas para ser cubiertos por los trabajadores.

Para efectos de nuestra investigación manejaremos la primera definición, debido a que es la definición que manejan los economistas en sus teorías de mercado de trabajo.

### **Demanda del mercado laboral**

“La demanda del mercado laboral son los puestos que demandan las empresas y que deben ser cubiertos por trabajadores que se adecuen a los requisitos exigidos por

ella. La demanda la constituyen la totalidad de puestos de trabajo vacantes, que son ofrecidos por las empresas a los trabajadores”. (Fontana, 2003; 122)

La demanda agregada del mercado de trabajo se constituye por las demandas individuales de empleo de cada una de las personas que desean obtener un puesto de trabajo. Esta demanda del mercado de trabajo es creciente con el salario ofertado por los empleadores. (Calvo, 2005; 39) . Al igual que en el caso de la definición de oferta del mercado de trabajo, podemos notar que la definición de demanda de mercado de trabajo de Fontana y Calvo es contraria, la primera hace referencia a que la demanda laboral se constituye por los puestos de trabajo que demanda las empresas y que deben ser cubiertos por trabajadores; la segunda en cambio afirma que la demanda del mercado laboral se constituye por el número de personas que desean trabajar a cambio de un salario. Al igual que en el caso anterior manejaremos la primera definición.

## Las variables

- **V1. Competencias laborales.-** «Una construcción a partir de una combinación de recursos (conocimientos, saber hacer, cualidades o aptitudes, y recursos del ambiente (relaciones, documentos, informaciones y otros) que son movilizados para lograr un desempeño». (LeBoterf, 1998)

## Dimensiones

- **V1.1. Competencias básicas.-** Son aquellas que se desarrollan principalmente en la educación inicial y que comprenden aquellos conocimientos y habilidades que permiten progresar en el ciclo educativo e integrarse a la sociedad. Tradicionalmente se incluyen entre las competencias básicas las habilidades en las áreas de lenguaje y comunicación, aplicación numérica, solución de problemas, interacción con otros y manejo creciente de tecnologías de información.
- **V1.2. Competencias conductuales.-** Son aquellas habilidades y conductas que explican desempeños superiores o destacados en el mundo del trabajo y que generalmente se verbalizan en términos de atributos o rasgos personales, como es el caso de la orientación al logro, la pro actividad, la rigurosidad, la flexibilidad, la innovación, etc.
- **V1.3. Competencias funcionales.-** Denominadas frecuentemente competencias técnicas, son aquellas requeridas para desempeñar las actividades que componen una función laboral, según los estándares y la calidad establecidos por la empresa y/o por el sector productivo correspondiente.
- **V2. Demanda Laboral.-** “La demanda laboral son los puestos que demandan las empresas y que deben ser cubiertos por trabajadores que se adecuen a los requisi-

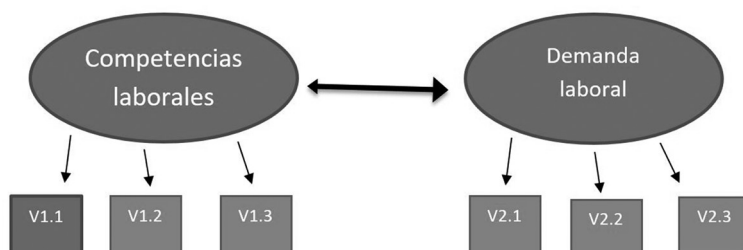
tos exigidos por ella. La demanda la constituyen la totalidad de puestos de trabajo vacantes, que son ofrecidos por las empresas a los trabajadores”. (Fontana, 2003; 122)

## Dimensiones

- **V2.1.-Perfil laboral.-** Capacidad efectiva para llevar a cabo exitosamente una actividad laboral plenamente identificada
- **V2.2.-Sueldo y salario.-** Es la retribución que los empleados reciben a cambio de su labor, es el elemento que permite, a la empresa, atraer y retener los recursos humanos que necesita, y al empleado, satisfacer sus necesidades materiales, de seguridad y de ego o estatus”.
- **V2.3.-Puesto ofrecido.-** Un puesto de trabajo es aquel espacio en el que un individuo desarrolla su actividad labor, también se le conoce como estación o lugar de trabajo

Gráfico 1

### Estructura gráfica de la competencia laboral y la demanda laboral



*Elaboración propia.*

El presente esquema representa gráficamente las posibles asociaciones entre la variable, competencias Laborales y Demanda Laboral. Es la presente investigación se ha dimensionado en tres categorías cada variable. Estas dimensiones serán cuantificadas variables través de sus indicadores, especificadas en la matriz de paralización de variables. Para cada dimensión se definirá al menos 5 indicadores. La escala de medición será ordinal.

## Metodología

La proyectada investigación será de tipo Aplicado, se analizará la información de forma descriptiva y se obtendrán indicadores del grado de asociación entre las competencias laborales y demanda laboral de los egresados y alumnos de los últimos años de la escuela Profesional de Ingeniería Económica-UNI, que vienen adquiriendo y su relación

con las necesidades de los sectores económicos de Lima metropolitana. El carácter descriptivo de la investigación se reflejará en la caracterización de los egresados de dicha Escuela, así como en la identificación de sus principales actitudes, saberes, conocimientos y la interrelación de los mismos con el campo laboral. Tendrá carácter analítico debido a que se trabajará con una serie de procedimientos para la recolección y síntesis de la información, a su vez se buscará la correlación de una serie de variables entre grupos determinados (egresados y estudiantes).

## Nivel de investigación

Es una investigación no experimental, correlacional y de corte longitudinal. La medición de las variables se realizará a través del instrumento de hoja de preguntas durante 20 días.

## Población

Los estudiantes egresados de los últimos cinco años la FIEECS \_ EPIEC. Dado que nuestro estudio se enfoca medir la relación entre la competencia laboral y la demanda laboral. Se tendrá en cuenta a los egresados que están laborando en las empresas públicas y/o privadas para tomar los datos de información primaria. La unidad muestral de análisis será el egresado entre los años 2010 y 2015.

Tabla 1

### Población de estudio

Año	Población/Estrato
2010	65
2011	66
2012	60
2013	61
2014	64
2015	67
Total	383

Fuente: ORCE

## Muestra

### Muestreo estratificado

Si se toma en cuenta los años de egreso de los estudiantes de la FIEEC\_EPIES, de la Universidad Nacional de Ingeniería, el cual cuentan con distintos tamaño en cada estrato. do y se aplicara el muestreo aleatorio estratificado, se obtendrá el tamaño de muestra. Con afijación proporcional al tamaño de la población.



Dado que nuestro estudio se enfoca medir la relación entre la competencia laboral y la demanda laboral. Se tendrá en cuenta a los egresados que están laborando en las empresas públicas y/o privadas para tomar los datos de información primaria. La unidad muestral de análisis será el egresado entre los años 2010 y 2015.

### Tamaño de la muestra

Para determinar el tamaño muestral y seleccionar los elementos de la muestra mediante un muestreo estratificado, a quienes se aplicaran los cuestionarios dentro de las unidades estadísticas elegidas El tamaño de muestra debe ser consistente con el objetivo general en nuestro caso determinar mediante la correlación de Pearson, el grado de relación entre las variables en estudio, y desde la distribución de una transformación de este el coeficiente se extrae una fórmula para determinar el valor global para “n”, el tamaño muestral:

$$n = \left\{ \frac{Z_{1-\alpha/2} + Z_{1-\beta}}{\frac{1}{2} \ln \frac{1+r}{1-r}} \right\}^2 + 3$$

Asumiendo un nivel de significación del 5%, una potencia de prueba 0.80, y un valor inicial de la correlación de 0.38, para aceptar la existencia de una relación entre las variables, se tiene que

$$n = \left[ \frac{1.96 + 1.05}{0.5 \times \ln \frac{1+0.88}{1-0.88}} \right]^2 + 3 = 59.6$$

Aplicando las correcciones por poblaciones finitas se tiene que el tamaño corregido es

$$n = \frac{59.6}{1 - \frac{59.6}{888}} = 171.5$$

El valor del tamaño muestral sera de n= 172 este valor sera distribuido según los tamaños de los estratos, usando la siguiente formula de afijación proporcional al tamaño del estrato

$$K = (172/383) = 0.44908616$$

Esto es para el calculo de los tamaños de muestra, se considera alrededor del 44.9% del tamaño de cada estrato, con la aplicación de esta valor proporcional en cada estrato, los resultados se presentan en la siguiente tabla estadística de distribución del tamaño muestral global.

Tabla 2  
**Tamaño de la muestra**

Año	Población	Muestra
2010	65	29
2011	66	30
2012	60	27
2013	61	27
2014	64	29
2015	67	30
<b>Total</b>	<b>383</b>	<b>172</b>

*Fuente : Elaboracion propia*

## **Recolección de la información**

### **Fuentes de información primaria**

La principal fuente de información serán los egresados de la EPIEC-UNI, mediante un cuestionario de preguntas. También se tendrá en cuenta las empresas donde laboran o laboraron y gerentes de recursos humanos de diferentes empresas relacionadas al campo económico-financiero-productivo. . Se implementara una serie de estrategias que permitan recopilar información pertinente de fuentes concretas. Dentro de las estrategias se sugieren:

- Establecer criterios de evaluación para la búsqueda de la información.
- Identificar la relevancia y la pertinencia de la información sobre el tema en cuestión.
- Eliminar la información que no sea coherente y que no tenga relación con la investigación.
- La información de la investigación será tomada de libros y revista académicas reconocidos.
- Los autores tomados de referencia deberán ser académicos y/o teóricos reconocidos.

### **Procesamiento de la información**

El procesamiento de la información respecto a esta investigación se realizara con criterios establecidos por el investigador en base a los objetivos que persigue la presente investigación. Para el procesamiento de la información se utilizaran software estadísticos como SPSS 19, Excel y en general el paquete de Microsoft Office. Los cuales servirán para sintetizar la información de encuestas y entrevistas en tablas y cuadros estadísticos.

## Métodos e instrumentos de obtención de datos

Si se tiene en cuenta que el método es medio, camino o estrategia concreta e integral de trabajo para el análisis de un problema o cuestión coherente con la definición teórica del mismo y con los objetivos de la investigación y La Técnica es el mecanismo que utiliza el investigador para recolectar y registrar la información tales como; Formularios, las pruebas psicológicas, las escalas de opinión y actitudes, etc.

En esta investigación se utilizará encuestas personales o entrevistas directas puesto que son las más usadas porque el entrevistador puede ofrecer una información más detallada, puede presentar gráficas, figuras, fotos, mapas, etc., además el entrevistador tiene la facilidad para resolver las dudas que tenga el entrevistado; sin embargo, el inconveniente es su alto costo, también se utilizara Entrevistas telefónicas puesto que muchas unidades de medición o elementos que conforman la población ( egresados de la FIEECS) se encuentran laborando en lugares fuera de la ciudad de lima lógicamente se tiene la tiene la limitante de poder entrevistar solamente a quienes tienen teléfono únicamente, no se pueden emplear ayudas audiovisuales y es más fácil que se rechace la encuesta; sin embargo son mucho más baratas que las entrevistas directas, además de producir grandes sesgos en la información obtenida.

Por otra parte la **Encuestas por correo**, su costo es relativamente bajo, permite la utilización de ayudas, fotos, figuras, mapas; da la oportunidad de volver a intentar con quienes no contestan. Sus desventajas son que no se puede controlar el proceso de respuesta, el tiempo que se toma el entrevistado y el orden en que se responde el formulario. No se puede desarrollar un proceso iterativo, no se aclaran dudas, no se garantiza que el entrevistado conteste las preguntas en el orden en el que se formularon, la eficiencia en nuestros países es un grave problema. **Grupos Focales o experimentos de laboratorio**, en ellos se reúne un grupo de personas en un lugar previamente definido y se les aplica la encuesta, tiene las ventajas de poder experimentar con el grupo, se pueden hacer modificaciones en presencia del grupo si así es conveniente y se puede procesar la información en presencia del grupo; su desventaja estriba en que difícilmente se pueda reunir a un grupo representativo de la población. Por la característica la presentes investigación todos estos métodos han sido recurso importantes para tomar decisiones y obtener el instrumento adecuado de medican de las variables en estudio en la presente investigación

## Sesgos en las respuestas

Shultz *et al.*, (1991), señala que cuando se aplican las encuestas o entrevistas, las respuestas de las personas pueden estar influidas por sesgos de diversa índole que afectan los resultados, por lo que es recomendable reducir dichos problemas para garantizar la confiabilidad y seguridad estadística al concluir el estudio.

Se distinguen dos grandes categorías: Sesgos instrumentales y No instrumentales; los primeros son de carácter operativo y que dependen en gran medida de la forma cómo esté estructurada la encuesta, mientras que los segundos, son los más difíciles de resolver, pues dependen de la actitud que tienen los entrevistados hacia la encuesta.

### **Sesgos instrumentales**

- **Sesgo del punto de partida:** aparece cuando en las preguntas iterativas, la cantidad inicial sugerida condiciona la respuesta final. Este sesgo se puede determinar si al subdividir la población en subgrupos e indicar a cada subgrupo una cantidad diferente, se obtienen resultados diferentes. Se puede eliminar obligando al encuestado a elegir la cantidad a pagar desde el inicio.
- **Sesgo de la información:** Ocurre cuando no se le informa a las personas la situación real del problema que se intenta valorar y tampoco del cambio que se propone; así que responden a la pregunta con poca certeza si con la cantidad indicada se logrará el cambio. Mitchell y Carson (1989), sugieren que no se debe llegar a los extremos, porque la descripción de escenarios de forma ineficiente o muy exagerada desde el punto de vista realista y objetivo, puede causar sesgos importantes en los encuestados. Es responsabilidad de los investigadores el proveer información científica, objetiva y profesional tanto como fuera posible.
- **Sesgo del investigador:** Cuando se realizan encuestas de forma individual, algunas personas dan respuestas exageradas, ante el temor de quedar frente al entrevistador como poco solidarias o inconscientes del problema planteado o sencillamente para querer agradar (Azqueta, 1994); Riera (1994), indica que la forma de abordar al entrevistado y las actitudes del entrevistados pueden tener un efecto significativo sobre la calidad de las respuestas obtenidas durante el proceso de valoración, también De la Maza (1996), hace ver que de la forma como se presenta quien encuesta o entrevista, da la partida para una buena o mala entrevista o incluso puede ser causa de rechazo.

### **Sesgos no instrumentales**

- **Sesgo de la hipótesis:** Aparece cuando el entrevistado no tiene ningún incentivo en brindar una respuesta correcta, por el hecho que la pregunta está basada en una situación hipotética.
- **Sesgo estratégico:** representa el sesgo más problemático para los economistas, ya que se presenta cuando la persona entrevistada no da una respuesta sincera y honesta, ya que asume que con su declaración puede influir en la decisión final

que se dé sobre la pregunta realizada, da una respuesta estratégica, es decir, la respuesta es manipulada según su conveniencia y la respuesta es falsa; este sesgo puede ser evitado a través de un diseño cuidadoso de la encuesta, para generar respuestas objetivas, viable y dependiente de muchas opiniones particulares.

## Instrumento de la encuesta

Con la orientación de expertos y la conceptualización del problema en estudio, se consideraron.

### Instrumento para medir la opinión de los egresados de la FIEECS

- **Bloque1.-** Se realiza la interrogante al entrevistado en relación datos generales de su formación profesional. También en esta parte se interroga a los entrevistados sobre las actividades profesionales.
- **Bloque 2.-** Se interroga al entrevistado sobre la variable Competencias Profesionales, su objeto era recolectar información sobre los indicadores de la Dimensiones: Competencias Básicas, Competencias Conductuales y Competencias funcionales.

El instrumento de la encuesta en este bloque los indicadores toman valores en la escala Likert de 1 a 5, en este caso el entrevistado dará su opinión subjetiva sujeto a su experiencia y vivencia personal en su formación profesional.

### Instrumento para medir la opinión de los empleadores en las empresas públicas o privadas

Se realiza la interrogante al entrevistado en relación datos generales de formación profesional.

Se interroga al entrevistado sobre la variable **Demanda laboral**, su objeto era recolectar información sobre los indicadores de la Dimensiones: Perfil Profesional, Sueldos y Puesto Ofrecido.

El instrumento de la encuesta en este bloque los indicadores toman valores en la escala Likert de 1 a 5, en este caso el entrevistado dará su opinión sujeto a su experiencia y vivencia personal en el desempeño profesional del profesional egresado de la FIEECS.

## Validación del instrumento

Para la obtención de información, referido a **egresados** de la Facultad de Ingeniería Económica, Estadística y Ciencias Sociales de la Especialidad de Ingeniería Económica

Tabla 3  
Estadísticos de confiabilidad

Reliability Statistics		
Cronbach's Alpha <sup>a</sup>	Cronbach's Alpha Based on Standardized Items <sup>a</sup>	N of Items
-,808	-,849	15

Fuente: Programa informático

El indicador de Alpha de Cronbach es -0.808, el cual nos indica la confiabilidad de Instrumentos es buena, puesto que esta por encima del mínimo que es 0.7

Tabla 4  
Estadísticos

Item Statistics			
	Mean	Std. Deviation	N
a1	3,0500	,88704	20
a2	2,6000	,82078	20
a3	2,9000	,71818	20
a4	3,1000	,78807	20
a5	3,0500	,68633	20
a7	2,9000	,85224	20
a8	2,7000	,65695	20
a9	2,8500	,81273	20
a10	3,0000	,91766	20
a11	3,2500	,78640	20
a12	3,0500	,82558	20
a13	2,9500	,88704	20
a14	2,9500	,88704	20
a15	3,2000	,83351	20
a16	3,0500	,88704	20

Fuente: Programa informático

Tabla 5  
**Estadísticas de ítems resumidos**

Summary Item Statistics							
	Mean	Minimum	Maximum	Range	Maximum / Minimum	Variance	N of Items
Item Means	2,973	2,600	3,250	,650	1,250	,029	15
Item Variances	,672	,432	,842	,411	1,951	,015	15

Fuente: Programa informático

Validación del instrumento para la obtención de información, referido a **Empleados** de los egresados de la Facultad de Ingeniería Económica, Estadística y Ciencias Sociales de la Especialidad de Ingeniería Económica

Tabla 6  
**Estadísticos de confiabilidad**

Reliability Statistics		
Cronbach's Alpha <sup>a</sup>	Cronbach's Alpha Based on Standardized Items <sup>a</sup>	N of Items
-,784	-,7008	20

Fuente: Programa informático

El indicador de Alpha de Cronbach es -0.808, el cual nos indica la confiabilidad de Instrumentos es buena, puesto que está por encima del mínimo que es 0.7

Tabla 7  
**Estadísticos**

Item Statistics			
	Mean	Std. Deviation	N
b1	3,3000	1,12858	20
b2	3,6500	,98809	20
b3	3,2500	1,06992	20
b4	3,8000	1,00525	20
b5	4,0000	,91766	20
b6	3,5000	,94591	20
b7	2,8500	1,13671	20

b8	3,6000	1,18766	20
b9	3,1000	1,07115	20
b11	3,5000	1,19208	20
b12	3,8000	1,00525	20
b13	3,4500	1,19097	20
b14	3,8500	,98809	20
b15	3,6000	1,14248	20
b16	3,5000	1,10024	20
b17	3,7500	1,01955	20
b18	3,5000	1,14708	20
b19	4,1500	1,08942	20
b20	3,4000	1,14248	20
b21	3,3500	1,03999	20

*Fuente: Programa informático*

## Procesamiento y análisis de datos

Dado que la investigación es del tipo correlacional, se encontrará las correlaciones entre las variables en estudio, las mismas que se presenta en la tabla 7.

A primera, una vista rápida de las principales correlaciones, indican que las principales variables están coleccionadas con magnitudes que van de 0.033 hasta 0.225, no todas son significativas teniendo en cuenta el nivel de significación del 5% .

- La variable Competencias Básicas y su relación con la variable Perfil profesional están asociados negativamente en forma significativa en un -20.10%
- La variable Competencias Básicas y su relación con la variable Sueldos están asociados significativamente en un 7%
- La variable Competencias Básicas y su relación con la variable Puesto Ofrecido están asociados negativamente en forma significativa en un -22.5% .
- La variable Competencias Conductuales y su relación con la variable Perfil profesional están asociados significativamente en un 4.2%
- La variable Competencias Conductuales y su relación con la variable Sueldos están asociados en un 3.3%
- La variable Competencias Conductuales y su relación con la variable Puesto Ofrecido están asociados en un 3.6%
- La variable Competencias funcionales y su relación con la variable Perfil profesional están asociados significativamente en un -7%
- La variable Competencias Funcionales y su relación con la variable Sueldos están asociados positivamente en un 16.2%



- La variable Competencias Funcionales y su relación con la variable Puesto Ofrecido están asociados positivamente en forma significativa en un 20.5%

Tabla 7  
Correlaciones de Spearman

			Correlations					
			Compe- tencias básicas	Compe- tencias conduc- tuales	Compe- tencias funcio- nales	Perfil profe- sional	Sueldos	Puesto ofreci- do
Spearman's rho	Competencias básicas	Correlation Coefficient	1,000	-,050	-,138	-,201**	,072	-,225*
		Sig. (2-tailed)	.	,607	,154	,037	,460	,028
		N	108	108	108	108	108	108
	Competencias conductuales	Correlation Coefficient	-,050	1,000	,029	,042	,033	,036
		Sig. (2-tailed)	,607	.	,769	,668	,734	,712
		N	108	108	108	108	108	108
	Competencias funcionales	Correlation Coefficient	-,138	,029	1,000	-,070	,162	,205*
		Sig. (2-tailed)	,154	,769	.	,474	,094	,038
		N	108	108	108	108	108	108
	Perfil profesional	Correlation Coefficient	-,201*	,042	-,070	1,000	-,064	-,025
		Sig. (2-tailed)	,037	,668	,474	.	,509	,795
		N	108	108	108	108	108	108
	Sueldos	Correlation Coefficient	,072	,033	,162	-,064	1,000	,004
		Sig. (2-tailed)	,460	,734	,094	,509	.	,967
		N	108	108	108	108	108	108
	Puesto ofrecido	Correlation Coefficient	-,225*	,036	,205*	-,025	,004	1,000
		Sig. (2-tailed)	,028	,712	,038	,795	,967	.
		N	108	108	108	108	108	108

\*. Correlation is significant at the 0.05 level (2-tailed).

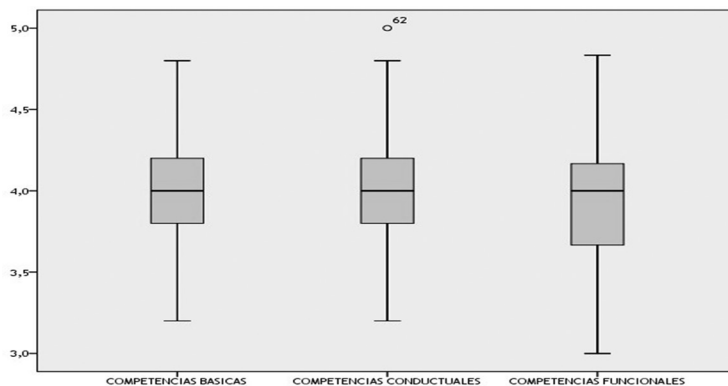
Fuente: Programa informático

### Diagrama de caja y bigotes

Las variables Competencias básicas y Competencias conductuales, relativamente están distribuidos en forma simétrica, con una mediana alrededor de 4, lo cual implica que el 50% de percepción respecto a estas competencias por los egresados están por encima de satisfacción buena. Mientras que las competencias funcionales su distribución es asimétrica, pero tiene una mediana alrededor de 4, pero con un sesgo de debajo de este valor, lo cual implica que hay muchos egresados que opinan dispersamente alrededor de regular referido a esta competencia.

Gráfico 2

**Distribución de competencias básicas, conductuales y funcionales**

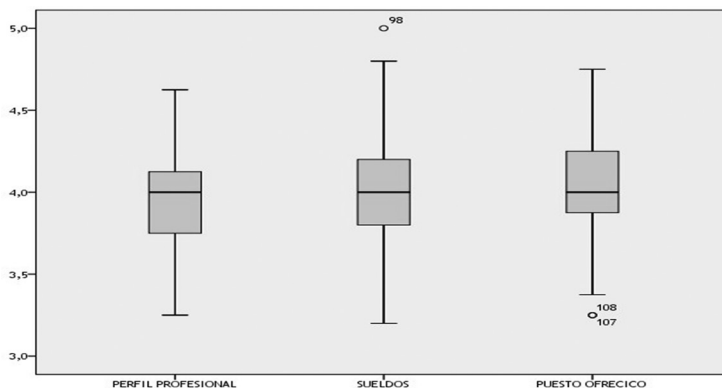


Fuente: Programa informático

La variable Competencias Sueldo relativamente está distribuida en forma simétrica, con una mediana alrededor de 4, lo cual implica que el 50% de percepción respecto a esta competencia por los Empresarios y/o Directivos están por encima de satisfacción buena. Mientras que las variables Perfil Profesional y Puesto Ofrecido su distribución es asimétrica, pero tiene una mediana alrededor de 4, pero con un sesgo de debajo de este valor, lo cual implica que hay muchos egresados que opinan nos dispersamente alrededor de regular referido a esta competencia.

Gráfico 3

**Distribución de perfil profesional, sueldos y puesto ofrecido**



Fuente: Programa informático

**Tabla 8**  
**Coefficiente de variabilidad**

Estadísticos descriptivos					
ITEMS		N	Media	Desv. típ.	CV
	Formación científica	108	3.90	.808	20.73%
	Formación Tecnológica	108	3.97	.767	19.30%
	Formación Humanística	108	3.89	.789	20.30%
Perfil profesional del egresado FIEECS	Aplicación de la Teoría Microeconómica	108	3.96	.808	20.39%
	Aplicación de la Teoría Macroeconómica	108	3.89	.813	20.90%
	Tiene habilidades de emprendedor	108	4.07	.770	18.90%
	Domina el inglés	108	4.04	.853	21.13%
	Domina relaciones humanas	108	3.92	.787	20.09%
	El sueldo está asociado a la labor que realiza	108	4.04	.808	20.01%
	El sueldo comparado con labores que realiza en otras instituciones	108	3.95	.813	20.57%
Sueldos determinados por...	El sueldo lo ha generado por su desarrollo personal. Capacitación	108	3.81	.775	20.32%
	El sueldo lo ha generado para satisfacer sus necesidades	108	4.15	.807	19.45%
	El sueldo porque se esfuerza para que la empresa crezca	108	3.93	.805	20.51%
Puesto ofrecido	Para Gerentes	108	4.13	.833	20.16%
	Subgerentes	108	4.06	.818	20.18%
	Subgerentes	108	4.09	.849	20.74%
	Jefe de proyectos	108	3.97	.803	20.20%
	Está relacionado a laborales asociado a su formación profesional	108	3.98	.809	20.31%
	Asistente	108	4.12	.840	20.38%
	Auxiliar	108	3.94	.812	20.63%
	Otros	108	3.93	.839	21.38%

Fuente: Programa informático

De la tabla del Coeficiente de Variación, sabemos que deberíamos tener un mayor Cv, ya que así evitamos outliers. Entonces podríamos trabajar nosotros con un Cv mayor a 19%, pero también perderíamos muchas variables debido a que todos están o tienen casi el mismo Cv. Esto es tanto para Bachiller, Magister y Doctorado.

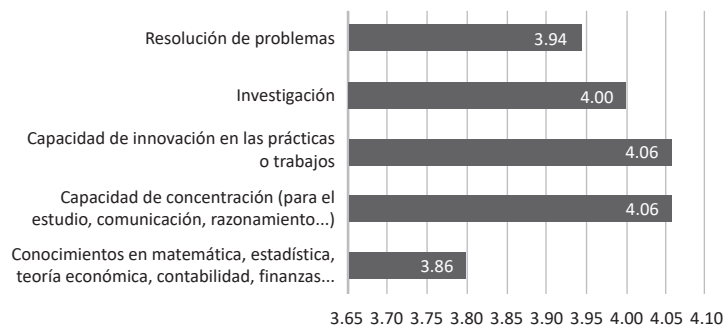
De la matriz de correlaciones, fijando un nivel de significancia de 0.05, observamos que solo algunas variables están correlacionadas, esto es porque son menores a  $\alpha=5\%$ . Asimismo, su correlación no es tan fuerte y es en algunos casos correlación negativa y en otros positiva. Por ejemplo, la variable Puesto para Subgerentes está correlacionado positivamente con la Formación Tecnológica que tenga el Egresado FIEECS. La variable Puesto para Auxiliar está correlacionado positivamente con la Formación Humanística del Egresado FIEECS.

Ahora veamos a continuación, las características más resaltantes de los egresados FIECS, de acuerdo a su grado académico.

## Bachiller

Gráfico 4

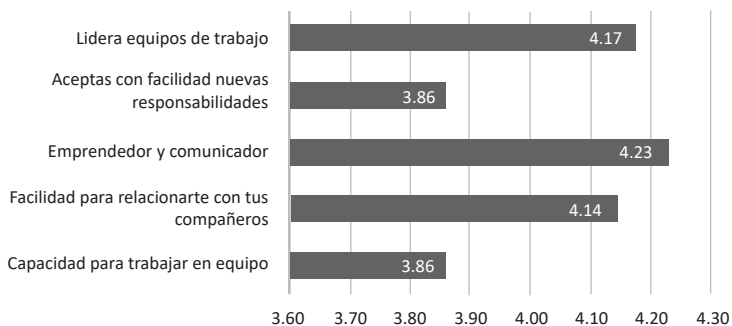
### Competencias básicas para Bachiller



Fuente: Elaboración propia

Las competencias básicas más resaltantes e importantes que el egresado FIECS valora con respecto a su paso por la EPIE-UNI son: Capacidad de concentración y Capacidad de innovación en las prácticas o trabajos.

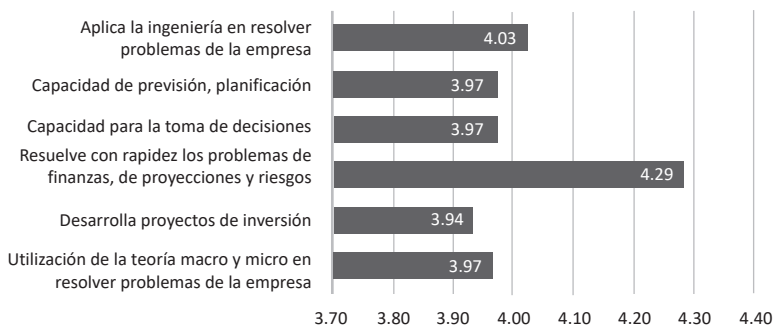
**Gráfico 5**  
**Competencias conductuales para Bachiller**



Fuente: *Elaboración propia*

En competencias conductuales, el Bachiller egresado de la FIECS, se caracteriza por liderar equipos de trabajo, ser emprendedor y comunicador, y tiene facilidad para relacionarse con sus compañeros muy satisfactoriamente.

**Gráfico 6**  
**Competencias funcionales para Bachiller**

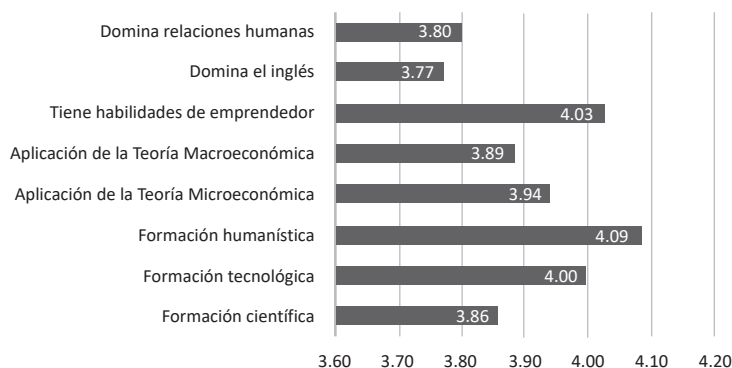


Fuente: *Elaboración propia*

En competencias funcionales, el Bachiller egresado de la FIECS, se caracteriza por resolver con rapidez los problemas de finanzas, proyectos de inversión, y aplicar la ingeniería para resolver problemas de la empresa muy satisfactoriamente.

Gráfico 7

**Perfil profesional para Bachiller**

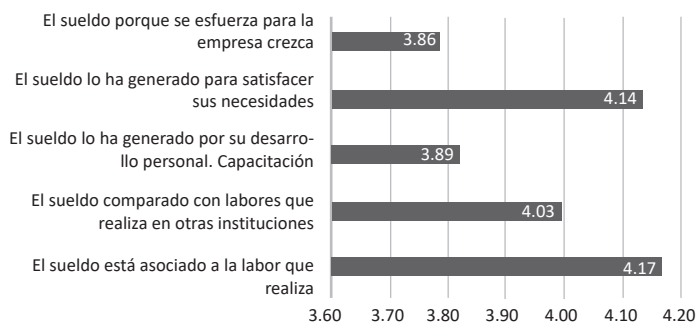


Fuente: Programa informático

Las apreciaciones más destacadas respecto al Bachiller, egresado de la FIECS son su Formación Humanística y tiene habilidades de emprendedor. Asimismo, cabe resaltar que el Bachiller egresado de la FIECS, domina de manera regular el inglés.

Gráfico 8

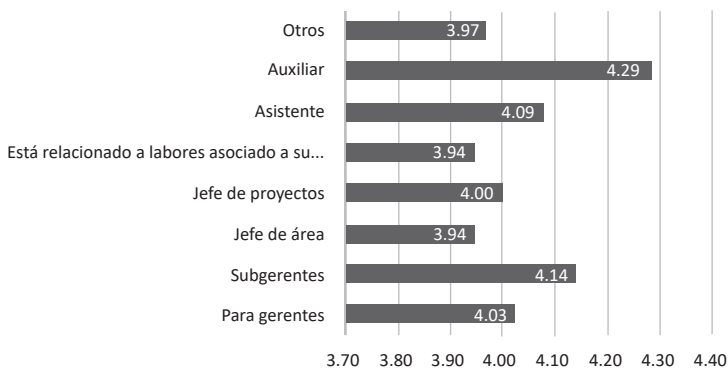
**Determinación de sueldos para Bachiller**



Fuente: Programa informático

Los sueldos determinados para el Bachiller egresado de la FIECS, se debería más a su labor que realiza, a la satisfacción de sus necesidades y con las labores que realiza en otras instituciones muy satisfactoriamente.

**Gráfico 9**  
**Puesto ofrecido para Bachiller**

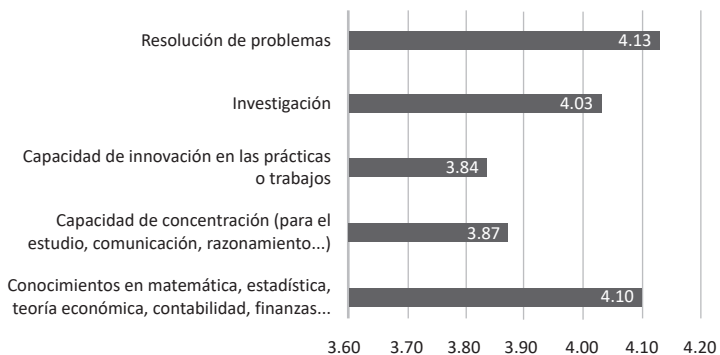


Fuente: Programa informático

En cuanto al puesto ofrecido para el Bachiller egresado de la FIECS, más resalta el cargo de Auxiliar, seguido de Subgerentes, Asistente y Gerentes muy satisfactoriamente.

## Magister

**Gráfico 10**  
**Competencias básicas para Magíster**

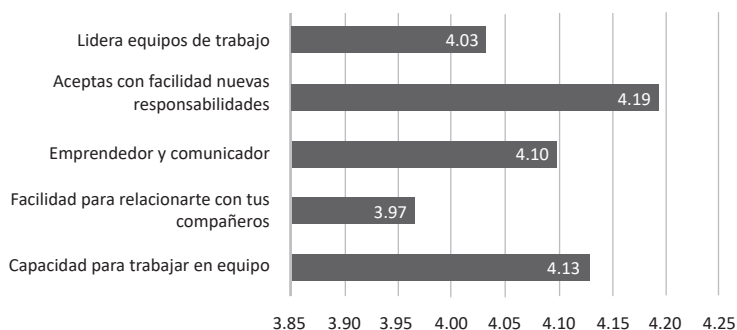


Fuente: Programa informático

Las competencias básicas más resaltantes e importantes que el Magister egresado de la FIECS valora muy satisfactoriamente con respecto a su paso por la EPIE-UNI son: Resolución de problemas, Conocimientos en matemática, estadística, teoría económica, contabilidad, finanzas, etc. e Investigación.

Gráfico 11

**Competencias conductuales para Magister**

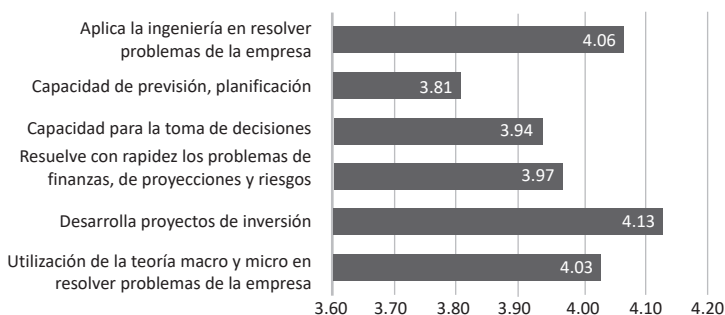


Fuente: Programa informático

En competencias conductuales, el Magister egresado de la FIECS, se caracteriza por Aceptar con facilidad nuevas responsabilidades, Capacidad para trabajar en equipo, Ser emprendedor y comunicador, y Liderar equipos de trabajo muy satisfactoriamente.

Gráfico 12

**Competencias funcionales para Magister**

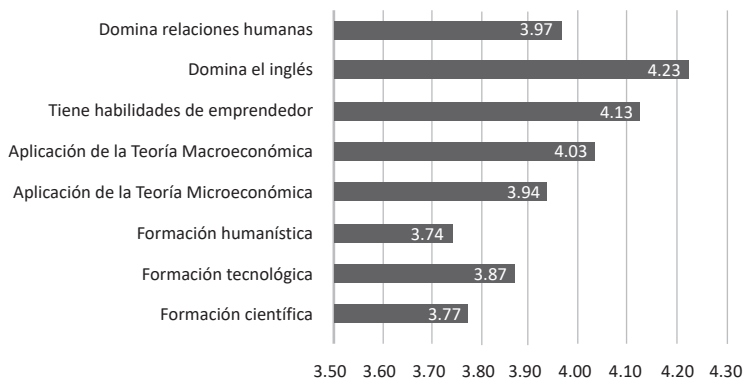


Fuente: Programa informático

En competencias funcionales, el Magister egresado de la FIECS, se caracteriza por Desarrollar proyectos de inversión, Aplicar la ingeniería para resolver problemas de la empresa, y Utilización de la teoría macro y micro en resolver problemas de la empresa muy satisfactoriamente.



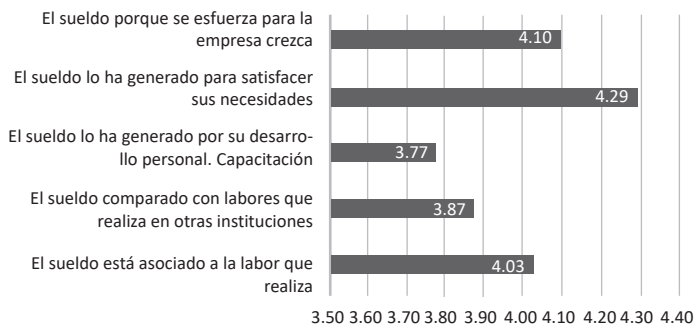
**Gráfico 13**  
**Perfil profesional para Magister**



Fuente: Programa informático

Las apreciaciones más destacadas de manera muy satisfactoria respecto al Magister egresado de la FIEECS son su: Dominio de inglés, Tiene habilidades de emprendedor, y Aplicación de la Teoría Macroeconómica.

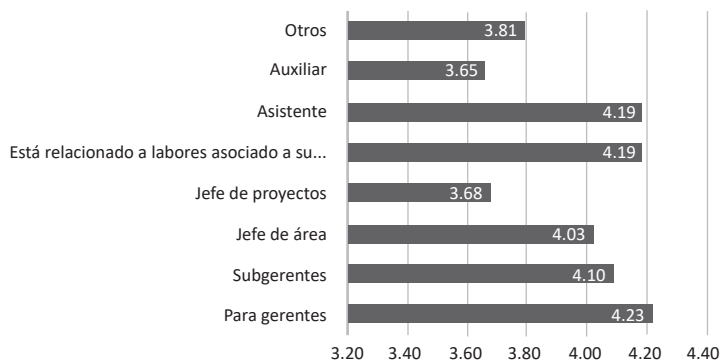
**Gráfico 14**  
**Determinación de sueldos para Magister**



Fuente: Programa informático

Los sueldos determinados para el Magister egresado de la FIEECS, se debería más a su: Satisfacción de sus necesidades, por su esfuerzo que realiza para que la empresa crezca y a labor que realiza muy satisfactoriamente.

Gráfico 15  
**Puesto ofrecido para Magister**

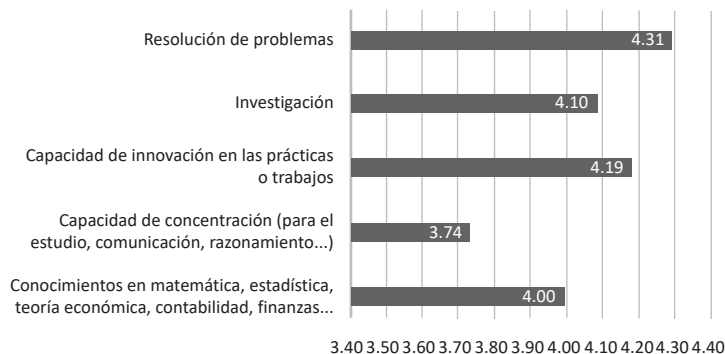


Fuente: Programa informático

En cuanto al puesto ofrecido para el Magister egresado de la FIEECS, más resalta el cargo de Gerentes, Asistentes, de Acorde a su formación profesional, Subgerentes y Jefes de área muy satisfactoriamente.

## Doctorado

Gráfico 16  
**Competencias básicas para Doctorado**

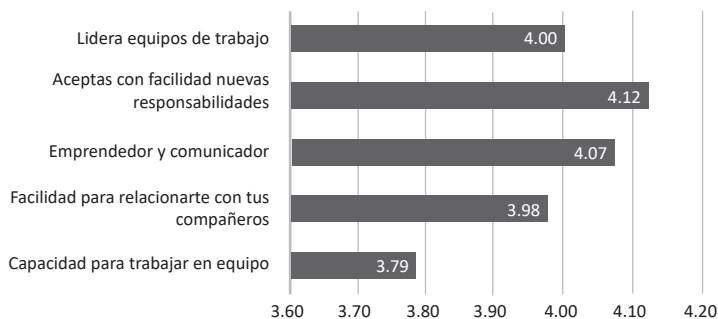


Fuente: Programa informático

Las competencias básicas más resaltantes e importantes que el Doctorado egresado de la FIEECS valora muy satisfactoriamente con respecto a su paso por la EPIE-UNI son: Resolución de problemas, Capacidad de Innovación en las prácticas o trabajos,

Investigación y Conocimientos en matemática, estadística, teoría económica, contabilidad, finanzas, etc.

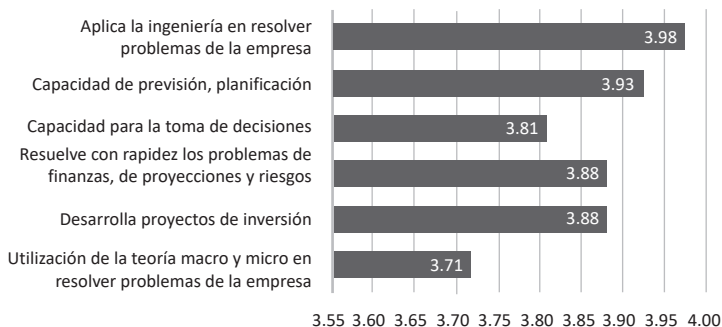
**Gráfico 17**  
**Competencias conductuales para Doctorado**



Fuente: Programa informático

En competencias conductuales, el Doctorado egresado de la FIECS, se caracteriza por Aceptar con facilidad nuevas responsabilidades, Ser emprendedor y comunicador, y Liderar equipos de trabajo muy satisfactoriamente.

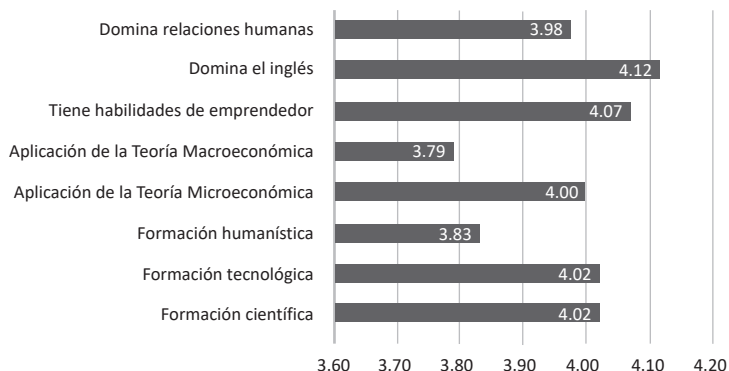
**Gráfico 18**  
**Competencias funcionales para Doctorado**



Fuente: Programa informático

En competencias funcionales, el Doctorado egresado de la FIECS, se caracteriza por Aplicar la ingeniería para resolver problemas de la empresa, Por su capacidad de previsión, planificación, Por desarrollar proyectos de inversión, y Resolver con rapidez los problemas de finanzas, de proyecciones y de riesgos satisfactoriamente.

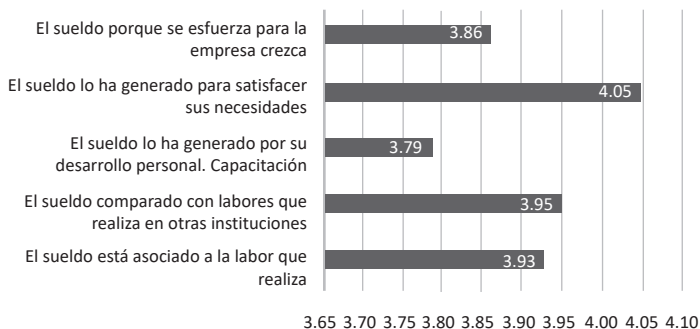
**Gráfico 19**  
**Perfil profesional para Doctorado**



Fuente: Programa informático

Las apreciaciones más destacadas de manera satisfactoria respecto al Doctorado egresado de la FIECS son su: Dominio de inglés, Tiene habilidades de emprendedor, Formación Tecnológica, Formación Científica, y Aplicación de la Teoría Microeconómica.

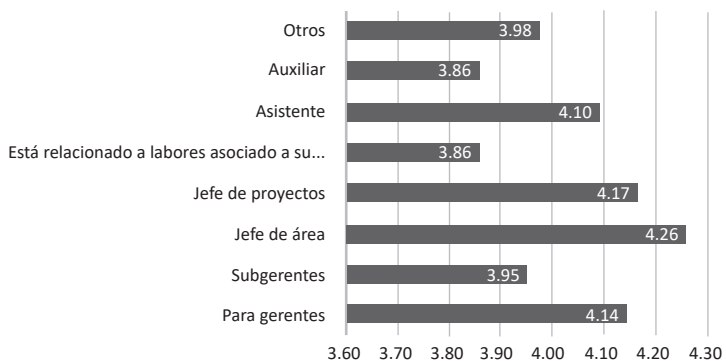
**Gráfico 20**  
**Determinación de sueldos para Doctorado**



Fuente: Programa informático

Los sueldos determinados para el Doctorado egresado de la FIECS, se debería más a su satisfacción de sus necesidades de manera satisfactoria.

Gráfico 21  
Puesto ofrecido para Doctorado



Fuente: Programa informático

En cuanto al puesto ofrecido para el Doctorado egresado de la FIECS, más resalta el cargo de Jefe de Área, Jefe de Proyectos, Gerentes y Asistentes muy satisfactoriamente.

## Conclusiones

Si se mejoran las competencias laborales de los egresados de la EPIEC UNI, estos tendrán una mejor inserción en el mercado laboral de Lima Metropolitana.

## Conclusiones específicas

- La determinación de las competencias laborales más solicitadas por los empleadores para estudiantes de economía se podrá actualizar la malla curricular de los estudiantes de la EPIE-UNI.
- El establecimiento de los requerimientos de los empleadores para estudiantes de economía se podrá afianzar la relación Empresa-Universidad-Egresados.
- El conocimiento de la relación Empleador -Egresados de la EPIE\_UNI las autoridades podrán obtener algunas recomendaciones para la mejora continua en dicha escuela.

## Recomendaciones

- Relacionar las competencias laborales y la demanda actual para egresados de Ingeniería Económica- UNI

- Proponer la actualización de la malla curricular de la EPIE basada en las competencias laborales requeridas por los empleadores
- Establecer una mejor relación entre EPIE-UNI empresa pública y privada
- Promover cursillos de capacitación docente te en competencias requeridas
- Sugerir algunos cambios en la práctica docente para favorecer los aprendizajes significativos
- Hacer partícipe al estudiante de su responsabilidad en la consecución de los objetivos de la calidad de la educación en la EPIE-UNI
- Dar algunas recomendaciones para el proceso de acreditación de la EPIE-UNI
- Fomentar la asociación de egresados de Ingeniería Económica- UNI.

## Referencias bibliográficas

Alonso, L.E., Fernández-Rodríguez, C. y Nyssen, JM<sup>a</sup> (2008). El debate sobre las competencias. Una investigación cualitativa en torno a la educación superior y el mercado de trabajo en España. ANECA, Madrid.

Álvarez R, P. (2004). Orientaciones al profesorado universitario para la enseñanza basada en competencias. VV. AA. Las competencias profesionales, mercado laboral y educación superior. La Coruña: Conferencia Internacional de Orientación, Inclusión Social y Desarrollo de Carrera

Ayats, J.C; Zamora, P. y Desantes, R. (2004). Los titulados de la Universidad Politécnica de Valencia y los empleadores. Servicio de Publicaciones de la UPV, Valencia.

Fundación Canaria Empresa Universidad de la Laguna (2007) ¿Qué demandan las empresas de los titulados universitarios. Fundación Empresa Universidad,

Universidad de la Laguna Bautista, et al (2012.) Un análisis de los profesionales en economía de la Universidad Militar de Nueva Granada: Mercado Laboral y Competencias. Revista Dianelt 02 de mayo del 2013

Benito, A y Cruz, A. (2005). Nuevas claves para la docencia universitaria. Madrid: Narcea.

Biggs, J. (2008). Calidad del aprendizaje universitario. Madrid: Narcea. Blanco Fernández, A. (2009). Desarrollo y evaluación de competencias en Educación Superior. Madrid: Narcea.

Castillo, S. y Cabrerizo, J. (2010). Evaluación educativa de aprendizajes y competencias. Madrid:

- Becerra, A. La Serna. (2014) Las competencias que demanda el mercado laboral de los profesionales del campo económico-empresarial en la actualidad. U del Pacífico
- Becker, G. El capital humano en la nueva sociedad
- Del Pozo, J. (2012). Competencias profesionales. Herramientas de evaluación: portafolios, rúbrica y las pruebas situacionales. Madrid: Narcea. 5 Flores O., R. (1999).
- Flores, H. (2007). Proyecto Tunning América Latina. Reflexiones y Perspectivas de la Educación superior en América Latina. Informe Final 2004-2007, universidad Deusto- Universidad de Groningen-Bilbao
- Fontana, L. Economía y Empresa. Universidad de Salamanca
- Kaluf, C. (2004). Competencias de egresados universitarios. CINDA. Chile Ley universitaria 30220
- Maldonado, M. (2012). Currículo con enfoque de competencias. Bogotá: ECOE. Mateo, J. (2000). La evaluación educativa, su práctica y otras metáforas. Barcelona: ICE Universitat.
- Ministerio de educación. (2009). Guía metodológica de evaluación de los aprendizajes en educación superior tecnológica. Lima: Impresos y sistemas. Ministerio de educación. (2014).
- Zabalza, M. (2003). Competencias docentes del profesorado universitario. Calidad y desarrollo profesional. Madrid: Narcea

# Crimen en Lima: una aproximación con datos distritales

Wilson Hernández Breña  
Universidad de Lima

## Resumen

Lima no solo mantiene niveles delictivos preocupantes sino también heterogéneos entre sus distritos. Pese a ello, son pocas las respuestas que se han dado a la pregunta más elemental: ¿qué causa el crimen en los distritos de Lima? Se usó el pool de datos de los siete años de la Encuesta Nacional de Programas Estratégicos (2010-2016) a fin de obtener artificialmente una muestra representativa de 35 distritos de Lima (N=53,787). Solo así fue posible responder a los dos objetivos de esta investigación: (1) analizar qué tan homogéneo es el crimen (y sus causas) entre un distrito y otro en Lima (análisis de clúster) y (2) identificar las razones que hacen que un distrito de Lima tenga más victimización que otro (modelación multinivel). Los resultados indican que, con el propósito de explicar las causas del crimen, es incorrecto tratar a Lima como un bloque homogéneo de distritos; estos, por el contrario, se pueden clasificar en tres grupos: protección latente, protección limitada y desprotección abierta, cada uno con una relación distinta con las tres teorías del crimen evaluadas (desorganización social, actividades rutinarias y capital social). Las implicancias apuntan a brindar mayor importancia a la multicausalidad del delito, mejorar y evaluar la participación local de la policía y contar con una mejor gestión de los incentivos económicos entregados a las municipalidades.

## Abstract

Lima suffers from a high crime rate, but one that is heterogeneously distributed throughout its districts. However, little is known about one of the basic questions regarding crime in the city: what causes crime among and across these districts? We constructed a data pool consisting of six years of data from the *Encuesta Nacional*

\* Esta investigación se desarrolló en el marco del programa de Investigación Formativa y contó con la participación de los alumnos Amós Canchari, Gustavo Chávez, Yordíño Jamanca, Astrid Mendoza, Álvaro Picón y Claudia Reyes.



*de Programas Estratégicos* (2010-2016) in order to obtain a representative sample for 35 districts in Lima (N=53,787). This allowed us to respond to the study's two main objectives: (1) analyze the extent of the heterogeneity of crime (and its causes) among Lima's districts (cluster analysis) and (2) identify the drivers that cause certain districts to have higher crime rates than others (multilevel modeling). Results show that we should not treat Lima as a homogenous city in terms of crime rate. Rather, we found that the city's districts could be classified into three groups (Latent Protection, Limited Protection and Permanent Defenselessness). We found that the theories of the origins of crime that we assessed in each group (social disorganization, routine activity theory, and social capital) differed in relation to the type of district. The policy implications of this research highlight the multicausality of crime, suggest improvements and assessments of police participation at the local level, as well as improving local management of economic incentives.

## Introducción

Pese a que el Perú es uno de los países con mayor incidencia delictiva en América Latina (Latinobarómetro, 2016), son pocas las respuestas (investigaciones) que tenemos a la pregunta más elemental y necesaria para formular políticas públicas efectivas: ¿qué causa el crimen?

A su vez, parte del estado pendiente de la respuesta a la pregunta se deriva de otra pregunta previa: ¿con qué data contamos? Desde el año 2010, la Encuesta Nacional de Programas Estratégicos (ENAPRES) ha producido datos sobre victimización delictiva en el Perú. Por esta fuente, sabemos que a nivel nacional una de cada cuatro personas (29%) ha sido víctima de algún hecho delictivo en los últimos doce meses (proporción con tendencia a la baja desde el 2010) y que solo un 14% de tales hechos son denunciados. En el 2015, el INEI realizó una encuesta a empresas en la que se preguntó por victimización, y aunque esta fuente ofrece otros resultados interesantes, en este estudio nos hemos concentrado en la victimización a personas.

El problema de fondo con la ENAPRES es su nivel de inferencia. Sus resultados son representativos a nivel nacional y regional, pero no a nivel distrital, lo que limita la capacidad de conocer cómo varía el crimen a este nivel. La data sobre victimización distrital es inexistente en el Perú, pese a que sería más provechoso identificar cuáles son las causas del crimen a este nivel de desagregación. De ser así, los resultados permitirían orientar no solo la política pública general (dando cuenta de macro relaciones) sino también la gestión municipal de la seguridad ciudadana (dando cuenta de micro relaciones y conexiones con la agenda pública); actualmente, esta se nutre de los datos de denuncias policiales que, debido a su subreporte, no reflejan la real dimensión del problema de inseguridad.

No contar con data distrital crea dos problemas prácticos. Primero, limitaciones para analizar la heterogeneidad distrital de la delincuencia en Lima (por ejemplo, solo

tres distritos están por debajo de la meta de 20% de victimización puesta por el actual gobierno para el 2021, mientras que otros 11 distritos tienen una victimización que supera el 30%). Segundo, limitaciones para identificar las causas sociales y económicas que hacen que la delincuencia sea más alta en un distrito limeño frente a otro.

No descartamos la ENAPRES como data para el análisis debido a que su uso estratégico permite brindar respuestas claras a la pregunta: “¿qué causa el crimen entre los distritos de Lima?”. A fin de superar las limitaciones advertidas, en este artículo construimos un pool de datos con las siete rondas de la ENAPRES (2010-2016) a fin de crear artificialmente una muestra representativa para 35 distritos de Lima (N=53,787). Son tres las razones por las que concentramos este trabajo en Lima. Primero, la estrategia de apilar datos favorece a aquellos distritos de mayor población, pues acumulan un mayor tamaño muestral con facilidad. Segundo, 54% de los cincuenta distritos con mayor población a nivel nacional pertenecen a Lima. Tercero, pese a sus diferencias, los distritos de Lima comparten cierta cultura local, lo que hace interesante estudiarlos en su conjunto.

Solo mediante estas operaciones pudimos luego responder las siguientes preguntas de investigación:

1. ¿Qué tan homogéneo es el crimen (y sus causas) entre un distrito y otro de Lima?
2. ¿Cuáles son las razones que hacen que un distrito de Lima tenga más victimización que otro?

El segundo paso estuvo en darle contenido teórico a las preguntas anteriores. El consenso es que son dos teorías las que explican la incidencia delictiva en unidades geográficas menores como los distritos: la teoría de la desorganización social y la teoría de las actividades rutinarias. Además, de la primera se deriva una teoría adicional que hemos independizado para evaluar su pertinencia y que corresponde al capital social. Esas tres teorías guían el resto del documento. Por último, una precisión: empleamos el concepto de crimen y no delito, pues lo que mide la ENAPRES son hechos delictivos que no siempre corresponden a delitos.

## **Marco teórico**

La falta de seguridad ciudadana es un problema que se ha vuelto más relevante en los últimos años. Al ser un determinante importante del bienestar de las personas, la carencia de seguridad genera una preocupación creciente en la sociedad, que exige al Estado aplicar políticas públicas eficaces. Diversas encuestas como el Latinobarómetro, concluyen que el Perú es uno de los países con mayor índice de victimización y que América Latina es la segunda región más violenta del mundo (Dammert y Arias, 2007).

La criminalidad es un fenómeno complejo y multicausal. Existen diversas teorías y enfoques que intentan explicar los determinantes del crimen. Nuestra propuesta enfoca las explicaciones del crimen en un nivel local (distrital) y, de esta forma, limi-

ta las teorías explicativas a un rango menor: desorganización social, capital social y actividades rutinarias. El enfoque de estas teorías pone un énfasis importante en las condiciones locales para la aparición del crimen.

Las ciudades mutan constantemente y operan en equilibrios subóptimos, por lo que su diseño y urbanización varían constantemente (Kubrin y Weitzer, 2003), y afectan las posibilidades de producción y control de la violencia. Para Wikström y Loeber (2000), dichas posibilidades se relacionan con las condiciones locales a fin de responder al por qué algunos distritos tienen más crímenes que otros, lo cual sugiere que distintas condiciones locales crean heterogeneidad en la respuesta frente al crimen.

Son dos las principales teorías que han buscado responder esta pregunta. La desorganización social y teoría de las actividades rutinarias. Ambas coinciden en el modo en que las condiciones locales que promueven o frenan la victimización se configuran en el espacio, sea porque algún factor clave está presente, ausente o contribuye a la degeneración de relaciones sociales. Para Smith et al (2000), la interrelación de la teorías de desorganización social y actividades rutinarias beneficiaría la comprensión de la aparición del crimen. A estas dos teorías se le suma una adicional, la teoría del capital social, que ha logrado un estatus de independencia dentro de la teoría de la desorganización social.

La desorganización social se define como la capacidad de una comunidad para supervisar y controlar el crimen mediante formas de control particulares basadas en redes locales (generadoras de capital social) y participación en organizaciones (Sampson & Groves, 1989).

La teoría de la desorganización social fue planteada inicialmente por Shaw y McKay (1942). La producción reciente alrededor de esta teoría es importante, pero fue el trabajo de Sampson y Groves (1989) el que la revitalizó. Esta teoría difiere de las teorías que explican el crimen a partir de las características individuales de las personas y se focaliza en el modo en que distintos barrios crean condiciones que favorecen o desfavorecen el crimen (Sampson y Groves, 1989). Surge como una forma de comprender la relación entre las características de los barrios y el crimen en zonas urbanas, entendiendo que tales características conllevan a espacios socialmente desorganizados (Kubrin y Weitzer, 2003). Para esta teoría, la forma de comportamiento colectivo de una comunidad determina el poder que tiene sobre su ambiente, que incluye la distribución espacial de la delincuencia.

La desorganización social aparece por el debilitamiento de las redes que mantienen el control social (Kubrin & Weitzer, 2003). El quiebre del control puede tener diversos factores disruptores que afectan la organización y derivan en violencia, como las características socio-espaciales de las ciudades y su heterogeneidad, la dificultad para absorber mano de obra de baja calificación o la provisión desigual de servicios públicos a la población (Muggah, 2012). Diversos estudios han dado cuenta del efecto negativo que ejerce la desorganización social sobre diversas formas de crimen, incluyendo incluso los homicidios (Lanier y Huff-Corzine, 2006).

El énfasis de la desorganización social está puesto en la estructura social de las zonas que analiza. Zonas con bajos niveles de desorganización social enfrentan dificultades permanentes para controlar el desorden social y el crimen. Kubrin y Wo (2015) resaltaron que en este tipo de comunidades se despliegan formas de control social informal, como mecanismos de vigilancia e intervención directa en problemas para enfrentar a gente sospechosa.

El énfasis puesto por la teoría de la desorganización social en las condiciones del barrio, ciertas características de la comunidad y, en general, la estructura social de espacios pequeños provoca que su relación con el crimen sea estadísticamente posible, pero también variable, lo cual es signo del modo en que la heterogeneidad entre distritos, barrios y manzanas puede crear distintos grados de asociación con la desorganización social. En esa misma línea, Bruinsma et al (2013) compararon diversos modelos de desorganización en Brasil a partir del uso de variables sugeridas por diversos académicos reconocidos en desorganización social. Hallaron que solo la heterogeneidad étnica se relacionó en forma directa y consistente con mayor criminalidad. Las redes amicales locales y la participación en organizaciones tuvieron efectos negativos en la reducción del crimen, al igual que la densidad poblacional. Sin embargo, otras variables, como la movilidad y estabilidad residencial, y el nivel socioeconómico no tuvieron efectos significativos en sus modelos. También en Brasil, Nogueira, Adrensen y Fonseca (2016) utilizaron regresiones espaciales para explicar diversas formas de delitos violentos. Solo usaron variables medidas a nivel distrital. De sus resultados, el robo de vehículos se redujo ligeramente en zonas con casas propias y aumentó en escasa proporción mientras más basura acumulada haya en las calles. Esta misma variable tuvo un efecto significativo en el aumento de robos comerciales. Mientras que la tenencia de títulos de propiedad puede ser un indicativo de nivel socioeconómico, mayor capacidad de protección y un ambiente menos peligroso, la acumulación de basura normalmente se toma como signo de desorden social; es decir, se interpreta como una forma de descuido que sugiere un débil control social formal e informal de diversos aspectos de la vida diaria, incluido el crimen.

Sampson y Groves (1989) consideran que la desorganización social es el otro extremo de la organización social (aproximación al capital social), en la medida que ambas aluden al uso de redes informales para promover la capacidad de resolver problemas comunes. Ellos encontraron que las redes de amistad, la participación formal y voluntaria en organizaciones, así como la habilidad de la comunidad para supervisar y controlar a los jóvenes, explican gran parte de las características exógenas del crimen y la victimización. Para Putnam (1994), el capital social no nace de la propia organización sino de las normas, valores y redes sociales que regulan los vínculos de asociatividad. Vínculos de este tipo contribuyen a repensar la lógica de algunas medidas de protección tomadas en los barrios. Mientras que algunas son de corte orgánico en tanto necesitan mayor coordinación –como la instalación y funcionamiento de organizaciones pro seguridad–, otras medidas consumen poca coordinación –como

la instalación de cámaras o vigilantes en la cuadra—. Debido a que no toda forma de capital produce seguridad es importante observar esta relación desde el punto de vista de Bourdieu. Para él, el capital social tampoco está en la propia organización, sino en los recursos insertados en las redes (Bourdieu, 1986).

El capital social permite alcanzar mayor cooperación en las comunidades. La evidencia sugiere que a mayores niveles de capital social menores niveles de criminalidad. Pero este efecto se produce solo si las organizaciones captan más miembros, socializan a los de menor edad y si su frecuencia de reunión es significativa (Bellair, 1997). En otras palabras, funcionan cuando hay cohesión y capacidad de respuesta local frente a condiciones adversas como la pobreza, urbanización, industrialización, desindustrialización, inestabilidad residencial, heterogeneidad técnica, etc. (Markowitz, Bellair, Paul, Liska, & Liu, 2001).

Desde la perspectiva racional del crimen, la desorganización social enfatiza la influencia de los factores situacionales (aspectos del ambiente físico, como iluminación y la presencia o ausencia de cámaras de seguridad) sobre el crimen, pero también remarca la influencia de las oportunidades que tiene un delincuente para cometer delitos (Barkan, 2012). Toda esta literatura es tributaria de Becker (1968) quien, en una época en que el crimen era explicado por las teorías de la anomia, incompetencias psicológicas y factores hereditarios, optó expresamente por evitarlas y, en su lugar, proponer un análisis basado en la elección racional. Los delincuentes realizarían un balance del costo de cometer un delito versus las posibles ganancias de hacerlo, sumado a la probabilidad de ser capturado, condenado y encarcelado. En dicha racionalidad, el control social y la efectividad de medios locales de controlar el crimen formarían parte de la decisión del delincuente de cometer un delito o no cometerlo.

Los factores situacionales a los que alude la economía del crimen han sido mejor desarrollados desde otro ángulo: el de la teoría de las actividades rutinarias. Cohen y Felson (1979) señalaron que para que un delito ocurra deben converger tres factores: agresores motivados, objetivos (víctimas) adecuados y ausencia de guardianes (policías) con capacidad para detener un delito. Para los autores, la ausencia de cualquiera de estos factores sería suficiente para reducir las probabilidades de que un delito se cometa. Es de nuestro particular interés centrarnos en las capacidades de los guardianes, especialmente de la policía.

Existe un largo consenso acerca del efecto de contar con más policías para enfrenar el crimen. Pero la mayoría de esta evidencia se ha concentrado —especialmente en los últimos años— en analizar el efecto de reubicar policías en las conocidas zonas calientes o *hotspots* (puntos de la ciudad con mayor concentración de delitos). La evidencia muestra que relocalizar los recursos policiales en los *hotspots* por un periodo de tiempo corto extingue las zonas de robo de frecuencia media (Santos & Santos, 2015). El análisis geoespacial incluso ha identificado regularidades estadísticas. Por ejemplo, un resultado común es que aproximadamente el 50% de los delitos se concentra en el 5% de intersecciones. Lo que este resultado ilustra es que la violencia

puede variar mucho entre una esquina y otra, pero que pocas de ellas ofrecen las condiciones para que la delincuencia prolifere. De ahí que la data de representatividad regional o provincial tenga una utilidad limitada para orientar las políticas públicas en espacios focalizados. Sin embargo, el consenso de la relación entre más policías y menos crimen ha sido criticado por vincularse con efectos exógenos, como los ataques terroristas (Di Tella, Galiani y Schargrodsky, 2010) y la condicionalidad del efecto en casos donde el número de efectivos es bastante bajo (Kleck y Barnes, 2014)

La teoría de la desorganización social tampoco está libre de críticas. Según Kubrin y Wo (2015), esta teoría presenta dos dificultades centrales. En primer lugar, no tiene la suficiente claridad sobre el rol explicativo de algunas variables al relacionar las condiciones estructurales exógenas de un barrio con el crimen; es el caso de los controles informales, los lazos sociales, la eficacia colectiva y el capital social. Investigaciones en Estados Unidos encontraron que la desorganización social es un factor que solo explica parcialmente el crimen. El estudio de Kingston, Huizinga y Elliott (2009) evaluó una larga lista de factores estructurales del barrio y halló que solo la pobreza explicó las conductas violentas de adolescentes. El trabajo de Lee, Lee y Hoover (2013) exploró las causas del robo de vehículos, tema poco trabajado en la literatura internacional. De sus predictores de desorganización social, solo la concentración de migrantes no tuvo un efecto significativo. Fue la desventaja social la que tuvo el mayor efecto sobre el aumento de robo de vehículos, mientras que la estabilidad residencial tuvo un efecto opuesto y significativo.

A pesar de existir un afán crítico a la teoría de la desorganización social, la evidencia a su favor, incluso en contextos culturales diferentes, es constante. Yang y Hoffman (1998) hallaron que la frecuencia de delitos en adolescentes estaba influenciada por el ingreso de su comunidad, la densidad poblacional, la proporción de familias monoparentales (pues reducen el control social en el hogar) y la mayor presencia de pares involucrados en la delincuencia.

En el fondo, la crítica de Kubrin y Wo (2015) a la desorganización social se relaciona más con la forma en que ha sido operacionalizada que con el contenido teórico. Resultados distintos entre un estudio y otro han sido atribuidos a que, por ejemplo, cada estudio toma variables diferentes para medir el control social (Steenbeek & Hipp, 2011).

Superando esta crítica, que más bien puede ser signo de la riqueza de la teoría, lo que aún no queda claro son los mecanismos por los que las desventajas sociales producen crimen. Kingston, Huizinga y Elliott (2009) señalan que se conoce poco sobre los procesos sociales que operan en forma similar entre barrios con distintos tipos de desventajas. Precisiones de este tipo crean la necesidad de estudiar patrones entre barrios o distritos, y así evaluar la forma en que, como grupo, estos barrios transforman sus activos y pasivos en mayor o menor criminalidad.

La segunda crítica de Kubrin y Wo (2015) a la desorganización social apunta no tanto a los métodos de medición, sino a lo que ha dejado de medir. Para ellos esta

literatura ha dado poca atención a los factores que se encuentran en la cultura del barrio, los controles sociales formales y la política económica urbana. Aunque se trata de variables más difíciles de evaluar, son aspectos con respuesta pendiente.

En resumen, las tres teorías revisadas, cada una desde su propia perspectiva, buscan explicar la criminalidad desde las características de los espacios locales y de las relaciones que surgen o existen en ellos. La desorganización social enfatiza las condiciones de los barrios y la asocia con la incapacidad comunitaria de supervisar y controlar el crimen, lo que genera señales de debilidad que son interpretadas por los delincuentes como señales para actuar con mayor facilidad. La hipótesis que se deriva de esta teoría es que el crimen tendrá una asociación negativa con formas de desorganización social que expresen similares condiciones. Como tal, esta teoría es muy dependiente de las condiciones locales, lo que propicia que pueda ser explicativa solo bajo ciertos contextos locales.

Por su parte, la teoría de las actividades rutinarias establece que la conjunción de agresores motivados, objetivos (víctimas) adecuadas y ausencia de guardianes (policías) propicia la acción criminal; sin embargo, dicha relación se halla siempre en función del stock de policías. Al respecto, nuestra segunda hipótesis es que el efecto de contar con una mayor disponibilidad de recursos policiales (comisarías y efectivos) tiene un efecto menor sobre el crimen, sobre todo porque dichos aumentos no afectan las condiciones de despliegue y presencia policial en forma significativa.

Finalmente, existen varias formas de aproximarse a medir el capital social. Sea a través de valores, asociatividad, confianza o recursos que existen en las redes, el capital social contribuye a una menor presencia del crimen. En ese sentido, nuestra tercera hipótesis, basada en la disponibilidad de data para evaluarla, es que el capital social, evaluado como las decisiones de implementar medidas de seguridad (resultantes de coordinación) y la confianza en la policía, están débilmente asociadas a la presencia de crimen. La razón se basa en que la forma de medir la variable capital social en la ENAPRES se sostiene en procesos de poca coordinación (cámaras de seguridad, vigilante y tranqueras) o coordinación orgánica limitada (juntas vecinales y alarmas comunitarias).

En la medida en que las tres hipótesis se asientan en características que pueden variar entre un distrito y otro (y también entre zonas de un mismo distrito), es de esperar que la asociación del crimen con la desorganización social, el capital social y las actividades rutinarias varíe. En otras palabras, la diferencia (heterogeneidad) entre distritos causa posiblemente que estas relaciones no sean homogéneas si no que existan subgrupos de distritos donde el crimen reaccione en forma distinta a sus predictores.

La multicausalidad del crimen no se reduce a las tres teorías aquí evaluadas. Su explicación también ha sido desarrollada desde la renovación de la teoría general de la anomia (Agnew, 1992), así como también por otras que atribuyen la incidencia delictiva a las etapas de desarrollo delinucional ligadas con la edad (Moffitt y Caspi,



1999), a motivaciones utilitaristas y gratificaciones hedónicas (Le Blanc y Fréchette, 1989), a la presencia de factores de riesgo estáticos y dinámicos en la persona (Andrews y Bonta, 1994), vulnerabilidades propias a las víctimas y formas de resiliencia (Walklate, 2011), estilos y rutinas que las hacen más propensas al crimen, entre otras. Especialmente los factores explicativos que giran sobre la decisión de delinquir pueden tener un efecto importante sobre el hecho que un distrito acumule más delitos que otro. Intentar identificar todos los factores causantes del crimen es un ejercicio siempre parcial, pero al mismo tiempo útil para identificar aquello que desde cada teoría permite ser identificado como un factor asociado al crimen.

## Lectura del crimen en Lima, distrito por distrito

¿Qué tan distinto es el crimen entre un distrito y otro en Lima? ¿Los distritos con menor victimización son también los que tienen las cifras más altas en los factores que contribuyen a la aparición del crimen? ¿Es el crimen una característica homogénea en toda la capital o, más bien, la heterogeneidad del crimen es algo que define a sus distritos?

Responder estas preguntas implica llevar a cabo una serie de decisiones metodológicas que, en lo fundamental, residen en una única premisa: ¿es posible estudiar el crimen y las diversas variables que señalan sus causas revisando individualmente los datos de cada distrito o, más bien, es mejor agrupar a los distritos similares bajo ciertos criterios?

La respuesta pasó por tres tareas. Primero, crear la data distrital para Lima a partir de la ENAPRES. Luego, analizar descriptivamente la data distrito por distrito a partir de variables de victimización, desorganización social, capital social y actividades rutinarias. Finalmente, abordar la segunda parte de la pregunta del párrafo anterior mediante análisis de clúster.

## Data

Se empleó las últimas siete rondas de la ENAPRES para construir un pool de datos (2010-2016) a fin de alcanzar representatividad distrital. Se trabajó solo con los distritos de Lima. De los 43 distritos de esta circunscripción, se seleccionó los 35 que cumplieron dos requisitos: muestra mayor de 331 observaciones (a fin de asegurar un error muestral máximo de 5.5%) o un coeficiente de disimilitud (*unlikeability*) de la victimización menor de 0.45 (Tabla 1). El uso de ambos criterios permite fortalecer la elección de los distritos incluidos en el análisis, ya que al mismo tiempo nos permite controlar la idoneidad de la muestra y la dispersión de los datos. El coeficiente de disimilitud pertenece a lo que se denominan indicadores de variación cualitativa, que miden la dispersión en distribuciones nominales y ordinales, para las cuales el coeficiente de variación es menor coherente. El coeficiente de disimilitud es una de las medidas más comunes para variables cualitativas (Kaplan, 2009). Como no es posible identificar la



distancia real entre el valor de una variable categórica (ser víctima de delito o no serlo), este indicador propone la fórmula: Coeficiente de *unlikeability* =  $2 \times (\text{Probabilidad de Categoría A}) (1 - \text{Probabilidad de Categoría B})$ , que mide variación y que también se interpreta como la probabilidad que dos observaciones obtenidas aleatoriamente sean distintas (Kvålseth, 1995). No existen parámetros para establecer rangos óptimos del coeficiente de disimilitud; su determinación queda a criterio del investigador. Se eligió 0.45 a fin de conservar la mayor cantidad de distritos y alejarse del punto máximo permitido (0.50). Ambos parámetros aseguran una dispersión adecuada en la data y la factibilidad para responder econométricamente a las causas del crimen. La muestra total fue de 53,787 personas. Se dejó nueve distritos de lado (Cieneguilla, Chaclacayo, Pucusana, Punta Hermosa, Punta Negra, San Bartolo, San Luis, Santa María del Mar y Santa Rosa), los cuales apenas representan el 2.8% de la población de Lima.

Tabla 1  
Distritos de Lima incluidos en la muestra

Distrito	N	CU	Distrito	N	CU
San Juan de Lurigancho	6,328	0.43	Rímac	1,230	0.42
Ate	4,073	0.43	La Molina	1,003	0.35
Villa El Salvador	3,407	0.46	San Miguel	980	0.39
Villa María del Triunfo	3,377	0.43	Surquillo	869	0.38
San Martín de Porres	3,357	0.43	San Borja	684	0.33
Comas	2,749	0.41	Breña	680	0.41
San Juan de Miraflores	2,546	0.43	Pachacámac	671	0.39
Chorrillos	2,443	0.41	Lurín	655	0.42
Lima Cercado	2,327	0.39	Pueblo Libre	601	0.37
Santiago de Surco	2,259	0.37	Jesús María	599	0.37
Los Olivos	1,992	0.41	Miraflores	548	0.31
La Victoria	1,627	0.43	San Luis	493	0.39
Puente Piedra	1,487	0.40	San Isidro	463	0.27
Lurigancho	1,487	0.42	Lince	426	0.35
El Agustino	1,480	0.45	Ancón	404	0.43
Carabayllo	1,368	0.42	Barranco	352	0.36
Independencia	1,348	0.39	Magdalena	344	0.32
Santa Anita	1,335	0.43			

CU = Coeficiente de disimilitud.

Fuente: ENAPRES, 2010-2016. Elaboración propia.

Desde el año 2010, la ENAPRES ha medido en forma constante la ocurrencia de estos diez hechos delictivos: (1) robo a hogares, (2) robo de auto, (3) robo de autopartes, (4) robo de moto o mototaxi, (5) robo de bicicleta, (6) robo de dinero, cartera o celular, (7) amenazas e intimidaciones, (8) maltrato psicológico y/o físico de parte de un miembro del hogar, (9) ofensas sexuales y (10) secuestro. Con esta información fue posible construir una variable dependiente de victimización que tome el valor de 1 cuando la persona haya señalado haber sido víctima de alguno de los hechos delictivos arriba señalados, y 0 en el caso contrario.

## **El crimen, distrito por distrito, en Lima**

En este punto, presentamos el ejercicio de construir el estado de la criminalidad en Lima a partir del análisis de datos de cada distrito. El objetivo es evaluar si es posible identificar, distrito por distrito, patrones asociados a la propia victimización y a lo que, según las teorías revisadas, se ha identificado como factores de riesgo o protectores frente a la criminalidad.

La conclusión anticipada es que si bien los “nuevos datos distritales” son útiles para evaluar a cada distrito de Lima y poner a cada uno en el contexto (creación de rankings), este análisis se hace más complejo al cruzar el ranking de victimización con los rankings de las múltiples variables que causan el crimen. Esto da pie a pensar que el análisis distrito por distrito es crucial pero que al mismo tiempo requiere de otros análisis que identifiquen patrones (clústeres) de distritos similares entre sí que puedan ser analizados de manera conjunta. Ampliaremos este punto en la subsiguiente sección.

El análisis está basado en datos muestrales, los errores estadísticos propios de este tipo de datos no poblacionales impiden parcialmente hacer referencia a diferencias significativas entre los indicadores de victimización de un distrito y otro. Todos estos datos están sujetos al margen de error (5.5%), lo que significa que la ubicación de un distrito en el ranking podría variar si se tuviera datos de todo el universo. Por ello, el análisis presentado solo analiza si un distrito se halla ubicado consistentemente en la parte alta del ranking o en la parte baja.

## **Victimización**

La victimización en Lima presenta dos características resaltantes. De un lado es dispersa (heterogénea) y de otro lado se asienta en una relación baja entre victimización patrimonial (robo de auto, autopartes, moto, bicicleta, dinero, cartera o celular) y no patrimonial (amenazas e intimidaciones, maltrato físico y/o psicológico de parte de algún miembro del hogar, ofensas sexuales y secuestro). Es probable que esta heterogeneidad marque distintas condiciones y recursos con los que cada distrito y su población enfrentan o resisten al crimen, lo cual da pie para pensar que Lima no puede ser leída en su conjunto sino que es importante pensar en subgrupos de distritos que reaccionen en forma similar antes las teorías del crimen.

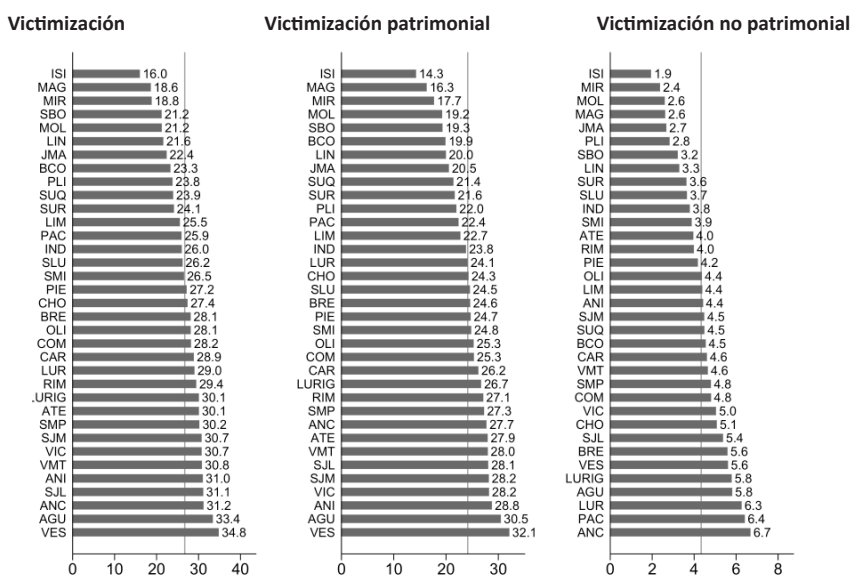
En Lima, la victimización tiene un grado de dispersión no despreciable. La dispersión (desviación estándar = 4.52) es solo ligeramente menor que la que existe entre regiones del Perú (5.20) y cercana a la que existe entre países de América Latina (6.02).

El ranking de victimización (en los últimos doce meses) de los distritos de Lima se presenta en la Figura 1. San Isidro es el distrito con menor porcentaje de personas víctimas de algún hecho delictivo (16%), seguido de cerca por Magdalena (18.6%). En el extremo opuesto se ubica Villa El Salvador con el doble de víctimas de la delincuencia (34.8%), y otros distritos como El Agustino, Ancón, San Juan de Lurigancho y Santa Anita. Es difícil establecer cuánto representa cualitativamente para una persona vivir en el primer grupo de distritos frente al segundo. Pero lo cierto es que la diferencia de la victimización entre San Isidro y Villa El Salvador es similar a la misma que existe entre la victimización en las regiones de San Martín (16%) y Puno (36%). Los distritos con más victimización no siempre son los que también presentan los mayores niveles de victimización no patrimonial.

Aun si se considera el margen de error muestral estimado para el pool de datos (5.5%), estos datos son bastante informativos. La variación de la ubicación de varios distritos entre un ranking y otro dificulta tener claridad sobre la concentración de la victimización. Si este análisis lo extendemos más allá de los cinco distritos en la parte alta y baja de los rankings, solo hará estos resultados sean menos legibles. Igual situación resulta al agregar y cruzar información de indicadores vinculados con las teorías que causan el crimen.

Figura 1.

**Victimización en últimos doce meses (promedio 2010-2016)**



Fuente: ENAPRES, 2010-2016.. Elaboración propia.

## Desorganización social

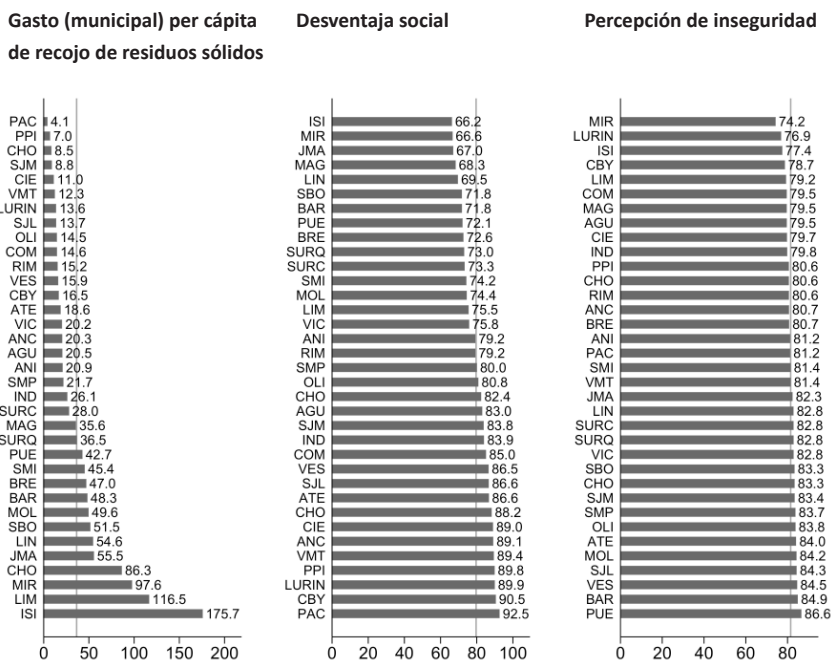
Son los distritos de mayor nivel socioeconómico los que presentan indicadores más bajos de desorganización social. Este resultado es consistente al margen que la desorganización social se mida como desventaja social, gasto (municipal) per cápita de recojo de residuos sólidos o percepción de inseguridad.

Existe una relación clara entre victimización y desorganización social, pero esta no siempre es absoluta. De un lado, distritos como El Agustino y Villa El Salvador son los que de manera consistente tienen los indicadores más altos de victimización total y patrimonial, pero no siempre los más negativos en términos de desorganización social (desventaja social y gasto municipal per cápita de recojo de residuos sólidos).

A diferencia de la victimización, la percepción de inseguridad muestra poca variabilidad entre un distrito y otro. Presenta una dispersión menor ( $DS=0.382$ ) y varía entre 74% y 86%. Su poca variabilidad es signo de los niveles de miedo y ansiedad por la delincuencia.

Figura 2.

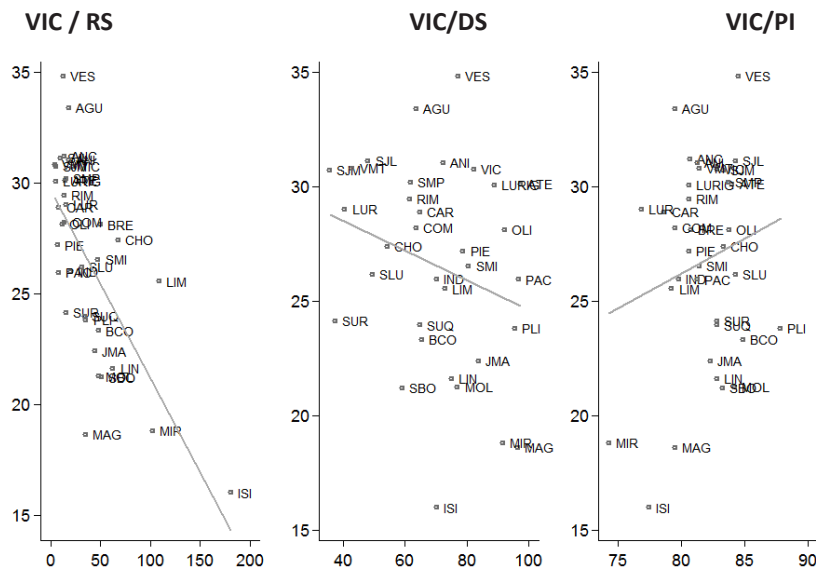
### Indicadores de desventaja social (promedio 2010-2016)



Fuentes: ENAPRES, Registro Nacional de Municipalidades.  
Elaboración propia.

Figura 3.

**Victimización en últimos doce meses (eje vertical) y desorganización social (promedio 2010-2016)**



VI: Victimización (eje vertical)  
 RS: Gasto (municipal) per cápita de recojo de residuos sólidos.  
 DS: Desventaja social.  
 PI: Percepción de inseguridad.

Fuentes: ENAPRES, Registro Nacional de Municipalidades. Elaboración propia.

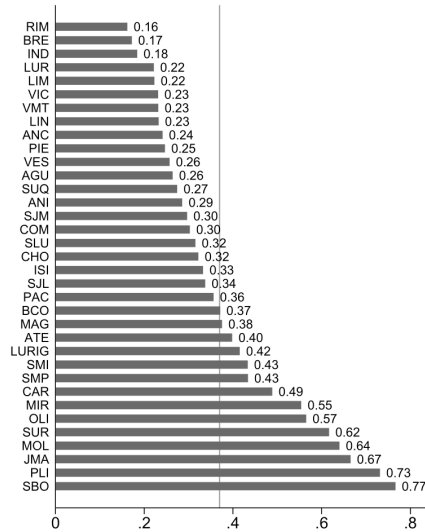
### Capital social

La mayoría de personas no ha implementado medidas en su barrio frente a la seguridad ciudadana (como instalación de cámaras, contratación de vigilante particular, instalación de alarma comunitaria, instalación de tranquera y constitución de junta vecinal); por ello, en la Figura 4 todos los valores distritales son menores que uno. Ello no esconde diferencias entre el comportamiento en cada distrito, pues mientras que las medidas son más comunes en San Borja, Pueblo Libre, Jesús María, La Molina y Surco, son hasta cinco veces menos frecuentes en los distritos que ocupan la parte baja del ranking: Rímac, Independencia, Lurín, Lima Cercado y La Victoria.

La relación entre victimización y capital social es negativa (Figura 5), aunque no muy fuerte. La debilidad de esta relación pone en evidencia una falta de homogeneidad en la distribución distrital de ambas variables en Lima. Por ejemplo, para niveles similares de victimización, Independencia y Pueblo Libre tienen datos muy distintos con respecto del número de medidas de seguridad tomadas en el barrio (0.18 y 0.73, respectivamente).

Figura 4.

**Número de medidas de seguridad tomadas en el barrio (promedio 2010-2016)**

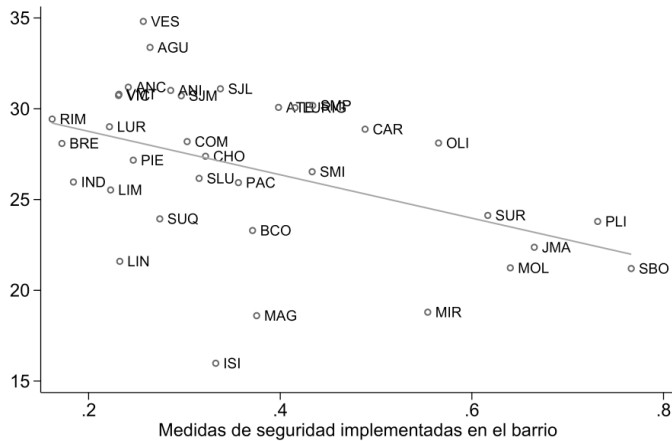


Fuentes: ENAPRES, Registro Nacional de Municipalidades. Elaboración propia.

Se preguntó por cinco medidas: instalación de cámaras, contratación de vigilante particular, instalación de alarma comunitaria, instalación de tranquera y constitución de junta vecinal.

Figura 5.

**Victimización (eje vertical) y número de medidas de seguridad tomadas en el barrio (promedio 2010-2016)**



Fuentes: ENAPRES, Registro Nacional de Municipalidades. Elaboración propia.

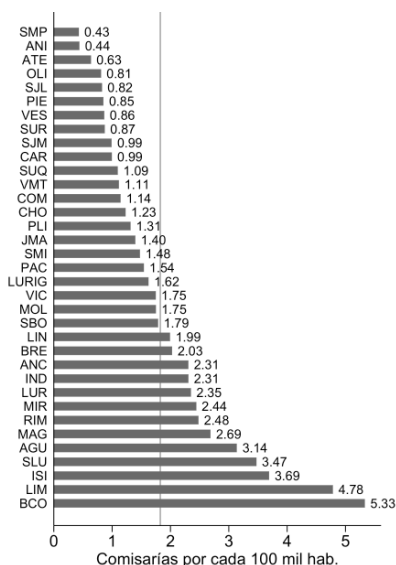
## Actividades rutinarias

Sucede lo mismo respecto de los recursos policiales y su relación con la victimización. En términos generales, a mayores recursos policiales, menor victimización (Figura 6). San Isidro es un caso particular: posee la prevalencia más baja de victimización en Lima e indicadores altos de disponibilidad de comisarías y policías, tiene 16% de victimización, y 3.69 comisarías y 398.9 policías por cada cien mil habitantes; sin embargo, otros distritos con más victimización, como Barranco y Lima Cercado, lo superan en número de comisarías por cada cien mil habitantes (5.33 y 4.78, respectivamente) y en términos de cantidad de policías por cada cien mil habitantes, Barranco lo supera con 404.4. En el otro extremo, los distritos con mayor victimización tienen menor disponibilidad de recursos policiales. Es el caso de Villa El Salvador, que tiene una alta victimización y una pobre disponibilidad de comisarías y de policías (0.86 y 58.5, respectivamente, por cada cien mil habitantes).

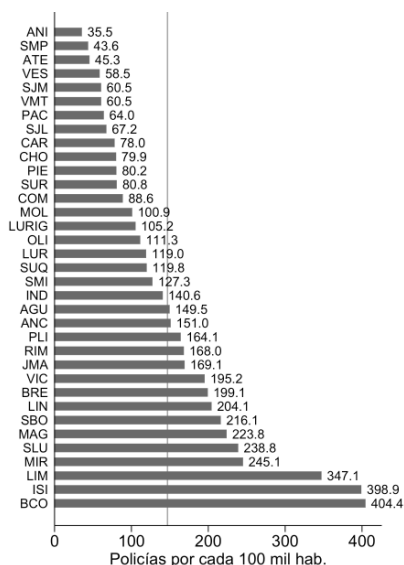
Figura 6.

### Comisarías y policías por cada 100 mil habitantes (promedio 2010-2016)

Comisarías por cada 100 mil hab.



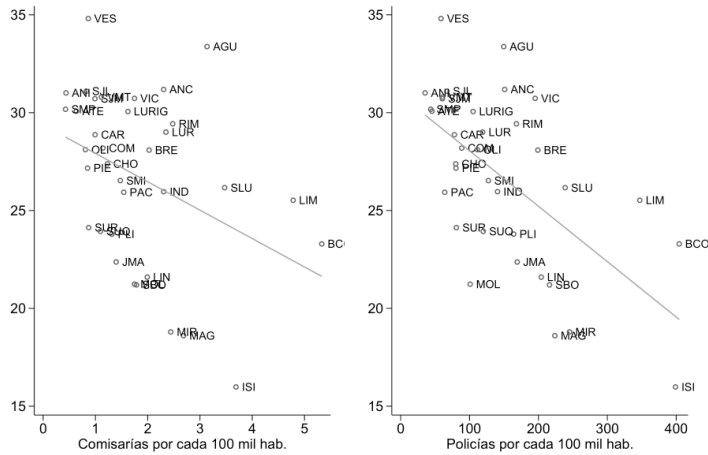
Policías por cada 100 mil hab.



Fuente: Censo Nacional de Comisarías. Elaboración propia.

Figura 7.

**Victimización (eje vertical) y comisarías y policías por cada 100 mil habitantes (promedio 2010-2016)**



Fuentes: Censo Nacional de Comisarías, ENAPRES.  
Elaboración propia.

**¿Es posible analizar la victimización distrito por distrito?**

A modo de cierre, los indicadores analizados líneas arriba nos permiten llegar a dos conclusiones centrales. En primer lugar, el crimen y los factores multicausales asociados a este tienen una importante variación entre un distrito y otro, pues se trata de factores que presentan una gran heterogeneidad entre los distritos de Lima. Y pese a que hay cierta recurrencia en distritos con una mejor o peor posición, no se trata de ubicaciones predecibles (están limitadas por el error muestral), y se han dado diversos ejemplos que quiebran dicha lógica.

En segundo lugar, el análisis distrital es productivo cuando se trata de pocas variables para unos pocos distritos. Sin embargo, este análisis se vuelve complejo cuando aumentan las variables por analizar y los distritos por comparar. Inclusive la idea de generar un ranking pierde solidez. Entender el crimen en 35 distritos de Lima es un ejercicio complejo, sobre todo si se desea describir y relacionar descriptivamente distintas formas de victimización con variables de desorganización social, capital social y actividades rutinarias.

La dificultad de basarse únicamente en este análisis se recoge de la mejor manera en la Figura 8. Ahí se compara el ranking de distritos en las variables de victimización, desorganización social, capital social y actividades rutinarias. En el escenario analíticamente ideal, Villa El Salvador —el distrito con la tasa de victimización más alta— debería ocupar el primer puesto en el resto de rankings ligados a victimización patrimonial y no patrimonial, desorganización social y actividades rutinarias. No obstante, este distrito no ocupa la misma posición en cada ranking. Nótese que este tipo de cruce

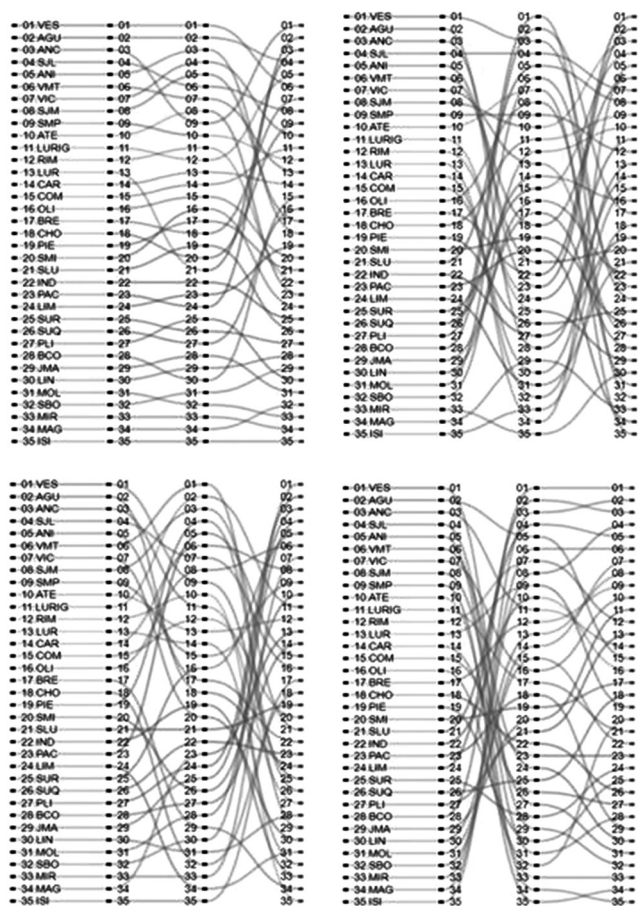


es mucho más acentuado en los distritos de la parte media de los rankings que se presentan en la Figura 8.

Como consecuencia de los dos puntos anteriores, es difícil estudiar la distribución distrital del crimen en Lima si no se recurre a técnicas que simplifiquen la diversidad de información. Una opción útil es aplicar técnicas de reducción de datos que permitan identificar a los distritos con mayor similitud (y a los diferentes entre grupo y grupo), y de ese modo hacer más comprensibles las relaciones que de ahí se obtuvieron. Ese es el objetivo de la siguiente sección, en la cual se aplicó análisis de clústeres.

Figura 8.

**Comparación de distritos según su puesto en variables clave**



VIC: Victimización. DS: Desventaja social. VICP: Victimización patrimonial. RS: Gasto (municipal) per cápita de recojo de residuos sólidos. VICNP: Victimización no patrimonial. COM: Comisarías por cada cien mil habitantes. PI: Percepción de inseguridad. POL: Policías por cada cien mil habitantes.

Fuente: ENAPRES, 2010-2016. Elaboración propia.

## Análisis de clústeres

### Metodología

La heterogeneidad de la victimización en Lima es importante. Mientras algunos distritos mantienen semejanzas entre sí, otros tienden a diferenciarse claramente. En otras palabras, hay subgrupos de distritos con los que, paralelamente, se evita el problema de trabajar con datos promedio y atomizar el análisis (enfoque distrito por distrito). El análisis de clúster es una de las técnicas que permitió identificar tales subgrupos y, de este modo, recuperar la estructura natural de la data.

El primer objetivo de la investigación se centra en describir la victimización en los distritos de Lima a partir de tres aspectos: estadísticas descriptivas y correlaciones entre las distintas formas de victimización y las variables arriba señaladas.

Se empleó el análisis de clústeres para identificar la existencia de patrones entre los 35 distritos del pool de datos. Este análisis permitió determinar las características que caracterizan a los grupos similares de distritos afectados por el crimen.

El análisis de clúster se realizó mediante el método no jerárquico *k-means*, el cual permite maximizar la homogeneidad dentro de los grupos de acuerdo con cierta cantidad de aglomeraciones o clústeres. Se empleó la distancia euclidiana al cuadrado de los centroides de cada clúster para medir proximidad.

En la literatura, no hay un protocolo único acerca del modo de elegir las variables que se emplearán para formar los clústeres. Las indicaciones se limitan a elegir las teóricamente relevantes, evitar que estén correlacionadas y seleccionar no muchas, ya que ello reduciría la diferencia entre clústeres (Mooi y Sarstedt, 2011). Por ello, se ha incluido las variables que cumplieron tres condiciones: vinculación con una de las teorías del crimen (Tabla 2), correlación no alta entre ellas (Tabla 3) y significancia al estimarlas en regresiones bivariadas con la variable dependiente (Tabla 4). En este último caso se empleó la base de datos colapsada al nivel distrital (35 observaciones).

Tabla 2.  
Variables a incluir en el cluster

Teoría	Variabes
Desorganización social	Percepción de inseguridad Gasto (municipal) per cápita de recojo de residuos sólidos
Capital social	Número de medidas de seguridad tomadas en el barrio
Actividades rutinarias	Comisaría por cada 100 mil habitantes

Tabla 3  
Correlaciones de las variables elegidas para construir los clústeres

	(1)	(2)	(3)	(4)
(1) Percepción de inseguridad	1			
(2) Número de medidas de seguridad tomadas	0.012***	1		
(3) Gasto (municipal) per cápita de recojo de residuos sólidos	-0.0161***	0.0078*	1	
(4) Comisarías por cada 100 mil habitantes	-0.031***	-0.057***	0.5824***	1

\*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$

Tabla 4.  
Estimaciones (mínimo cuadrados ordinarios) con variables elegidas para el cluster como independientes (variable dependiente: victimización)

	(1)	(2)	(3)	(4)
Percepción de inseguridad a delitos patrimoniales	0.304			
Número de medidas de seguridad tomadas		-58.17***		
Gasto (municipal) per cápita de recojo de residuos sólidos			-0.0849***	
Comisarías por cada 100 mil habitantes				-1.452**
Observaciones	35	35	35	35
R2	0.033	0.194	0.485	0.144

\*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$

La técnica del *k-means* es flexible en tanto permite al investigador elegir el número de clústeres que se formarán. Se testearon diversas soluciones (con 2, 3, 4, 5 y 6 clústeres). Se eligió construir tres clústeres. Esta elección pasó por tres decisiones. Primero, descartar la solución que produjo clústeres con pocos distritos. Segundo, validar los clústeres evaluando diferencias significativas mediante MANOVA (análisis multivariado de la varianza), la que permite evaluar la relevancia global de distintos modelos que asumen comportamientos distintos en cada clúster. Tercero, descartar las soluciones con más clústeres en las que no era clara la lógica de la agrupación.

El algoritmo de construcción del clúster identifica un punto central de inicio con base en la identificación aleatoria de un grupo de observaciones. Luego, a partir de dichas observaciones la técnica permite identificar gradualmente grupos cercanos en distancia. Esto implica que cada vez que se repite la rutina de programación para la creación de clústeres, la conformación de cada clúster (distritos) varía. En el caso de esta investigación, la variación fue menor. Normalmente, se observó que tres distritos cambiaban de clúster en la solución final; para evitar este tipo de inconve-

nientes y resultados no replicables se fijó como grupo de inicio un conjunto definido de observaciones.

La conformación de los clústeres se presenta en la Tabla 5. Es importante aclarar que los clústeres reflejan características similares entre cada grupo de distritos y características distintas entre clústeres. Si bien un clúster puede contener un distrito considerado parecido al de otro clúster, es importante aclarar que no proponemos una reconfiguración de Lima (equivalente a pensar en Lima Centro, Norte, Sur y Este).

Los clústeres implican que sus distritos comparten características similares en torno a la desorganización social, capital social y actividades rutinarias, las cuales los hacen distintos de los distritos de otro clúster. En la Tabla 5 se resume tales características pero, más importante aún, se presenta una propuesta de denominación para cada clúster que varía entre la protección y la desprotección, que será analizada en la siguiente sección: protección latente, protección limitada y desprotección abierta.

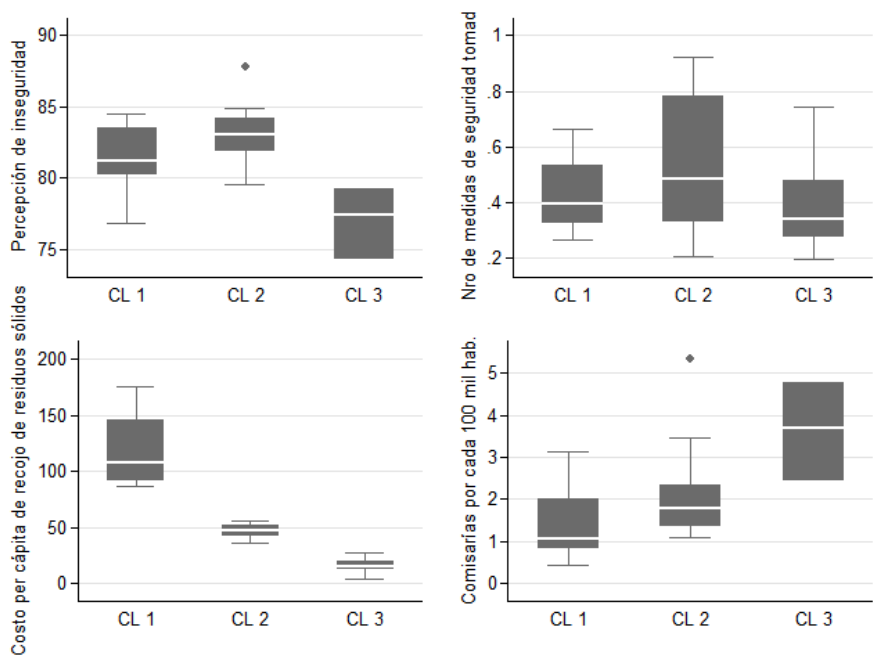
Tabla 5  
Conformación de clústeres

	Clúster 1	Clúster 2	Clúster 3
	Lima Cercado Miraflores San Isidro	Barranco Breña Chorrillos Jesús María La Molina Lince Magdalena Pueblo Libre San Borja San Miguel Surquillo	Ancón Ate Carabayllo Cieneguilla Comas El Agustino Independencia La Victoria Los Olivos Lurigancho Lurín Pachacámac Puente Piedra Rímac San Juan de Lurigancho San Juan de Miraflores San Martín de Porres Santa Anita Santiago de Surco Villa El Salvador Villa María del Triunfo
Victimización	Media	Media	Alta
Desorganización social	Baja	Media	Alta
Capital social	Medio	Alto	Bajo
Actividades rutinarias	Alta	Media	Baja
Nombre de clúster	Protección latente	Protección limitada	Desprotección abierta

*Elaboración propia.*

Figura 9.

Gráfico de caja para variables empleadas para la construcción de los clústeres



Fuente: ENAPRES, 2010-2016.

Elaboración propia.

## El crimen por clústeres en Lima

El análisis de clúster produjo tres clústeres de distritos que hemos denominado Protección Limitada (N=3,551), Protección Limitada (N=10,185) y Desprotección Abierta (N=45,297). Nótese que ningún clúster hace visible un estado de protección absoluto. Su denominación obedece a una protección siempre condicionada por factores terceros. Los tamaños de la muestra en cada clúster permiten contrastar estadísticas entre ellos a fin de identificar patrones, pero antes de realizar un análisis detallado es importante remarcar tres características generales de los clústeres identificados.

En primer lugar, cada clúster agrupa a distritos que, según intuición previa, son similares pero también incluye a otros que, según esa misma intuición, no deberían formar parte de un mismo grupo. Por ejemplo, el Clúster 1 está conformado por Lima Cercado, Miraflores y San Isidro. La diferencia de estos dos últimos distritos frente al primero es evidente, sin embargo, son varias las razones que le dan coherencia la conformación del Clúster 1. La construcción de los clústeres se realiza con base en un conjunto de variables que en la literatura especializada han probado ser criminógenas y que, de acuerdo al análisis de clúster, son más similares entre sí en comparación

con otros distritos. Además, en cada distrito estos factores se manifiestan en forma diversa: Miraflores tiene una gran presencia de turistas y público flotante, mientras que San Isidro es un distrito de población móvil, que esencialmente va al distrito para trabajar. Asimismo, el análisis de clúster admite cierta heterogeneidad entre sus observaciones. Si San Isidro se ubica en el Clúster 1 es porque presenta una menor distancia hacia las observaciones de los otros distritos del mismo clúster que a otros clústeres. En el mismo sentido, el Clúster 3 incluye a Surco, un distrito que normalmente no sería ubicado en ese grupo. La hipótesis es que Surco es un distrito con zonas de alta desigualdad y que además limita con distritos de su propio clúster (movilidad de la delincuencia).

En segundo lugar, existe cierto patrón en el orden de los clústeres –característica que hemos denominado consistencia global–, es decir, en la mayoría de casos el Clúster 3 (Desprotección abierta) tiene los valores negativos muy altos en variables que reflejan victimización, desorganización social, capital social y actividades rutinarias. Por ejemplo, en el Clúster 3 el porcentaje de victimización es mayor (29.9) y es el que tiene el menor número de comisarías por cada cien mil habitantes (1.6). La consistencia global es positiva en tanto las variables empleadas no presentaron una correlación importante. Además, en casi todos los casos existen diferencias visibles para cada variable entre un clúster y otro, aspecto que sugiere la heterogeneidad entre distritos (lo que refuerza la lógica de emplear análisis de clúster).

En tercer lugar, la consistencia global coexiste con lo que para efectos prácticos llamamos consistencia temática. Por ejemplo, el Clúster 1 suele ser siempre el que tiene los indicadores de victimización más bajos y lo mismo sucede con su posición en el resto de variables de desorganización social, capital social y actividades rutinarias.

Una vez entendida la lógica de los clústeres, se describirá cada uno utilizando los valores promedio obtenidos por cada clúster, respecto de las variables de victimización, desorganización social, capital social y actividades rutinarias.

### **Clúster 1: Protección latente**

AAgrupaa los distritos con la menor proporción de víctimas de hechos delictivos (20.1%), junto con bajos índices de desorganización social y niveles medios de capital social. Este estado se complementa con la mayor disponibilidad de recursos (policiales) para el control del delito. Este clúster comprende tres distritos: Lima Cercado, Miraflores y San Isidro. En conjunto, los distritos de Protección latente cuentan con una mezcla de factores de riesgo ligados con la desorganización social, pero también con los mayores recursos que poseen sus habitantes, lo cual los hace atractivos para ciertas formas de delincuencia.

Aun cuando es el que tiene la menor ocurrencia de victimización de los tres clústeres, solo llega a niveles medios (por ejemplo, su nivel de victimización es similar al promedio de victimización en América del Sur). En otras palabras, este clúster tiene

mayores niveles de protección que el resto de clústeres, pero dado el contexto de alta victimización y recursos siempre escasos, dicha protección no puede asumirse como continua sino como latente. El riesgo de ser víctima de la delincuencia siempre existe y solo permanece, aparentemente, inactivo. De ahí que el clúster haya sido denominado de Protección Latente, riesgo que se grafica bastante bien en el 77% de personas que cree que será víctima de algún delito (Figura 16).

Las diversas formas de extensión del crimen —como victimización patrimonial y no patrimonial; victimización indirecta y reiterada— son menores en este clúster, dado que es el que se encuentra mejor protegido, por ello tanto la victimización patrimonial como la no patrimonial son las de menor incidencia, 28.5% y 3.4%, respectivamente (Figura 11).

En términos de expansión delictiva, la victimización se ha desarrollado menos en este clúster porque existe un menor porcentaje de miembros del hogar (victimización indirecta) que han sido víctimas de la delincuencia (Figura 12). También es menor la victimización reiterada, es decir, la cantidad de veces que una persona fue víctima de algún delito en los últimos doce meses. Incluso los delitos violentos (cometidos con arma) tienen menor incidencia (20.1%) (Figura 13); pero en este caso la diferencia es menor frente al Clúster 2 (23.7%). Esta cercanía es la que hace que la protección en este clúster no sea absoluta sino relativa o latente.

Las condiciones de desorganización social en el Clúster 1 son menores en comparación con el resto de clústeres. Presentan el mayor gasto per cápita de recojo de residuos sólidos a nivel municipal, así como la mejor ubicación en el indicador de desventaja social (78.3).

El capital social no sigue una lógica gradual entre clústeres, como en el caso de la victimización. Detrás de este resultado parece haber una respuesta mayor que complementa la relativa tranquilidad del distrito con los recursos con los que cuentan sus habitantes y las medidas que, en ese escenario, pueden o deben tomar sus habitantes. El clúster de Protección latente es el segundo en importancia en cuanto a la aplicación de medidas de seguridad en el barrio. Incluso es el que menos medidas orgánicas (alarmas comunitarias y juntas vecinales) implementó (14.1%). En este clúster, el hecho de contar con mayores recursos policiales y menor victimización podría haber limitado la cooperación entre vecinos con fines de protegerse, cuando este tipo de protección se hubiera favorecido con los recursos privados de las propias personas.

Finalmente, visto desde la teoría de las actividades rutinarias, el Clúster 1 presenta los mejores indicadores de recursos policiales. Triplica tanto la disponibilidad de comisarías (3.6 versus 1.4 por cada cien mil habitantes) como la disponibilidad de policías del Clúster 3 (330 versus 95 por cada cien mil habitantes). Solo en este clúster se cumple con el estándar internacional de la Oficina para la Droga y el Delito de la ONU, que señala la disponibilidad 300 policías por cada cien mil habitantes.

En resumen, el clúster de Protección latente es el que ofrece a sus habitantes las mejores condiciones en un contexto de alta victimización y percepción de inseguri-

dad como el limeño; pero tales condiciones siempre están sujetas a los insuficientes recursos para controlar la delincuencia. De ahí que dicha protección tenga siempre la característica de ser latente —o de riesgo en apariencia inactivo— frente a los otros clústeres, donde lo que predomina en forma abierta es alguna forma de desprotección

## **Clúster 2: Protección limitada**

Es un clúster de tamaño medio. Incluye diez distritos: Barranco, Breña, Chorrillos, Jesús María, La Molina, Lince, Magdalena, Pueblo Libre, San Borja, San Miguel y Surquillo. Muchos de sus indicadores muestran valores negativos más altos que los del clúster de Protección latente, pero mejores que los del clúster de Desprotección abierta. En este clúster, el nivel de victimización es medio y se combina con una mayor extensión de diversas formas de victimización, generalmente asociadas a bienes patrimoniales (autos, motos, bicicletas) y la presencia de mayor victimización no patrimonial. El mayor nivel adquisitivo en este clúster y la mayor presencia de policías hacen que sus residentes dependan menos de formas orgánicas de capital social (alarmas comunitarias y juntas vecinales). El Estado y su control social existen pero funcionan con eficacia relativa. El tema de fondo es que al analizar los indicadores de victimización, se observa que no son suficientes para contrarrestar la criminalidad. De ahí que el nombre de este clúster aluda a una protección que, en efecto, existe, pero que tiende a ser limitada y es percibida como tal, como lo demuestra el hecho de tener el porcentaje más alto en inseguridad (83.1%).

La victimización en este clúster (23.7%) es más cercana a la del clúster anterior (20.1%) que al de Desprotección abierta (29.9%) (Figura 10). En general, el resto de características de este clúster también lo hacen más parecido al clúster de Protección latente; por ejemplo, la victimización patrimonial llega al 31.6% y la no patrimonial al 3.6% (Figura 11). Lo mismo ocurre con la victimización indirecta (familiares en el hogar víctima de la delincuencia) y la victimización reiterada (Figura 12).

Las condiciones que propician la desorganización social en el clúster de Protección limitada son relativas. El indicador de desventaja social fue el segundo entre los tres clústeres con 72.9 (Figura 15). El gasto de recojo de residuos sólidos se ubicó en valores medios (46.7 soles por habitante), pero muy lejano al del clúster de Protección latente (119 soles por habitante). La combinación de carencias sociales considerables y la poca preocupación de los gobiernos locales por reducir los riesgos que propician la aparición del crimen en el entorno urbano, crean un ambiente donde la protección está limitada, especialmente cuando se considera los factores que la población tiene para enfrentar a la delincuencia (capital social) y los recursos policiales para enfrentarla (teoría de las actividades rutinarias).

De un lado, el capital social tiene una presencia clara en el clúster de Protección limitada. Se ha identificado un patrón en apariencia no aleatorio en el número de medidas de seguridad adoptadas en el barrio. Es en este clúster donde más se ha adoptado estas medidas, especialmente las no orgánicas (vigilante, cámaras de vigilancia



y tranqueras). Del otro lado, los recursos policiales para el control social agrandan las carencias ya identificadas en el clúster anterior. Por cada cien mil habitantes, existen 2.1 comisarías y 187 policías (el estándar óptimo es de 300).

En resumen, los distritos de este clúster ofrecen protección. Pero se trata de una protección que no está garantizada desde la política pública de seguridad ciudadana con suficiente disponibilidad de comisarías y policías ni necesariamente con las condiciones estructurales (desorganización social) que hacen que el delito pueda ser menos frecuente. La protección es limitada y las formas de victimización se focalizan en aquellos recursos que aparentemente tienen menor capacidad de ser protegidos en la población de estos distritos.

### **Clúster 3: Desprotección abierta**

Es el clúster más grande. Agrupa a los distritos con los valores más negativos a lo largo de casi todos los indicadores analizados. No solo es el grupo de distritos con la más alta victimización (29.5%), sino que supera el promedio nacional. Sus niveles de desorganización social son marcadamente más altos frente al resto de clústeres, razón por la cual probablemente algunas mediciones de capital social son más importantes en este clúster ya que funcionan como contrapeso para aquello que el Estado le brinda en cantidad baja (escasa disponibilidad de comisarías y policías). En estos distritos, la presencia de otros problemas sociales es más común. De ahí que la lista de distritos en este clúster sea importante: Ancón, Ate, Carabayllo, Comas, El Agustino, Independencia, La Victoria, Los Olivos, Lurigancho, Lurín, Pachacámac, Puente Piedra, Rímac, San Juan de Lurigancho, San Juan de Miraflores, San Martín de Porres, Santa Anita, Santiago de Surco, Villa El Salvador y Villa María del Triunfo.

Los distritos clasificados en Desprotección abierta lideran el ranking de victimización en sus distintas formas de expansión. La victimización indirecta, aquella que contabiliza los miembros del hogar víctimas de algún delito en los últimos doce meses, es marcadamente más alta (64.8%) en este clúster que en el resto (Figura 12). Lo mismo sucede con la victimización repetida (Figura 12). Incluso el indicador de delito violento, cuando el delincuente usó arma, es marcadamente más alto en el clúster de Desprotección Abierta (29.5%) (Figura 13).

Por otro lado, de acuerdo con la teoría de las actividades rutinarias, la materialización del robo a hogares suele producirse cuando confluyen un delincuente motivado, un objetivo (víctima) adecuado y la falta de guardianes (Cohen y Felson, 1979). La menor presencia de recursos policiales en estos distritos reduce la probabilidad de una vigilancia óptima, lo cual, junto a menores recursos para proteger los hogares, crea una potencial vulnerabilidad que podría ser aprovechada por algunos delincuentes.

En el caso de los robos de dinero, cartera o celular, suelen ser formas de criminalidad menos organizadas ligadas a la delincuencia común; aspecto que sugiere que las condiciones de estos distritos y de sus habitantes hacen propicia la aparición de

estas formas no organizadas de delitos. Junto a ello, los mayores niveles de desorganización social hallados en el clúster de Desprotección abierta sugieren la existencia de mayores factores de riesgo criminógenos, los cuales elevan la probabilidad de que parte de la población caiga en la delincuencia. Un dato por tener en cuenta es que el 43.6 % de reos, cuya última residencia fue algún distrito de Lima, proviene de distritos de este clúster (Censo Nacional de Establecimientos Penitenciarios, 2016). Si quitamos a Lima Cercado, perteneciente al primer clúster, el porcentaje de reos del Clúster de Desprotección abierta pasa a ser 71.3%.

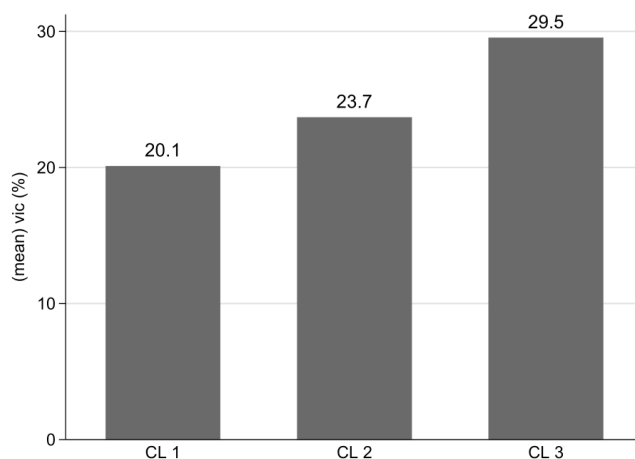
Este último dato se relaciona con otra de las grandes explicaciones del crimen: la desorganización social. La lectura de las siguientes cifras grafica condiciones estructurales que son poco atendidas por la gestión local y que –por ello mismo– adquieren un tono permanente. El indicador de desventaja social fue marcadamente negativo en este clúster (67.2) (Figura 15). Carencias de este tipo hacen alusión a necesidades estructurales no atendidas, cuyo cambio difícilmente se da en el corto plazo, y ello genera factores de riesgo criminógenos permanentes en el tiempo. Cuando tales factores de riesgo están condicionados por el ambiente (calles mal iluminadas, calles tomadas de urinarios, basurales, etc.), la gestión local puede atenderlos y reducir los puntos de riesgo que atraen delincuencia (Wilson y Kelling, 1982). Sin embargo, en el clúster de Desprotección abierta, la otra *proxy* para medir desorganización social muestra una imagen deteriorada de los distritos considerados. El gasto municipal de recojo de residuos sólidos es de 15.8 soles por habitante, cifra nueve veces menor a la del clúster de Protección latente (Figura 15).

En escenarios adversos como el descrito, aparecen balances. Una de estas vías de balance es la formación del capital social. El capital social es un activo importante en este clúster de distritos con Desprotección abierta. Las diferencias se captan cuando diferenciamos a las medidas de seguridad entre orgánicas (alarmas comunitarias y juntas vecinales) y no orgánicas (vigilante, cámaras de vigilancia y tranqueras). En el clúster de Desprotección abierta se recurre con mayor frecuencia a las primeras (Figura 14). Las medidas orgánicas incluyen la instalación de alarmas comunitarias y juntas vecinales; implican mayor organización continua en el tiempo; o buscan cooperar con las autoridades estatales o bien suplir su ausencia. Es la carencia de un Estado efectivo el que empuja a estas formas de capital. Las medidas no orgánicas de parte del Estado requieren no solo menos coordinación sino mayor capital económico (cámaras, tranqueras y vigilantes).

En línea con lo anterior, los recursos para el control social desde el Estado son pobres, pese a la magnitud de carencias y condiciones que propician la aparición de la criminalidad. La disponibilidad de comisarías y policías por cada cien mil habitantes es baja. Como se señaló líneas arriba, la Oficina para la Droga y el Delito de la ONU recomienda contar con 300 policías por cada cien mil habitantes y en este clúster, se cuenta apenas con un tercio de lo recomendado (95) (Figura 17). Brechas de este tipo son difíciles de sobrellevar en el corto plazo. Si bien es cierto que los residentes

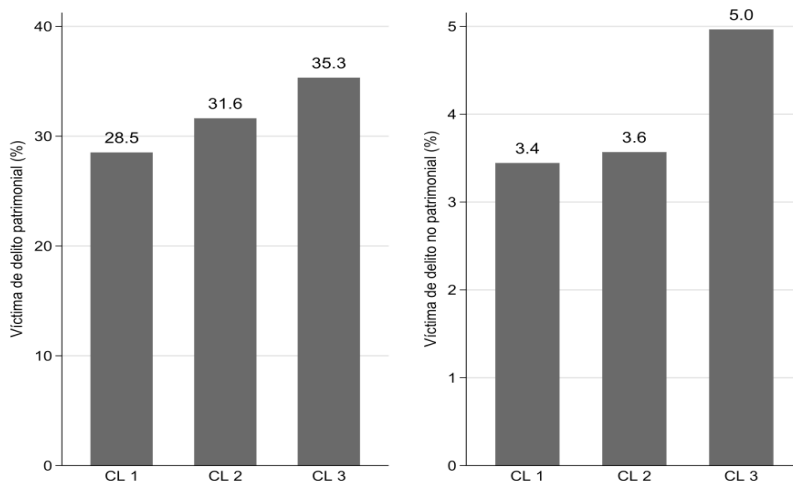
de los distritos de este clúster pueden hacer una vida normal, en términos analíticos viven en espacios de desprotección permanente, lo cual implica tener mayores probabilidades de ser víctima de algún delito.

Figura 10.  
**Victimización (en últimos doce meses) por clúster**



Fuente: ENAPRES, 2010-2016.  
Elaboración propia.

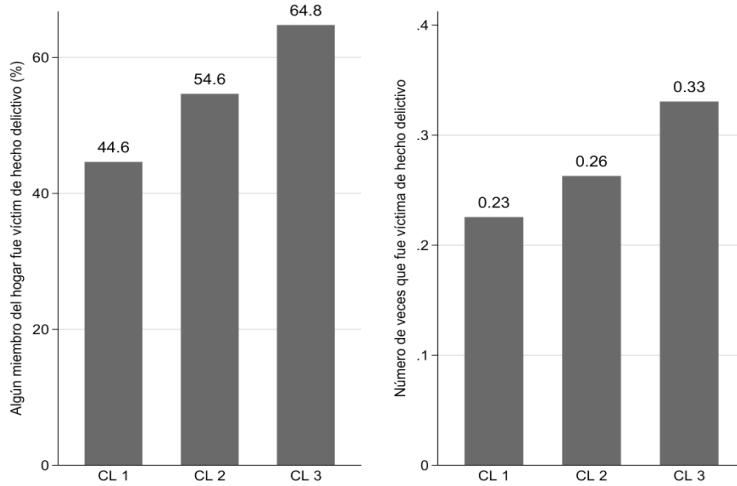
Figura 11.  
**Victimización patrimonial y no patrimonial (en últimos doce meses) por clúster**



Fuente: ENAPRES, 2010-2016. Elaboración propia.

Figura 12.

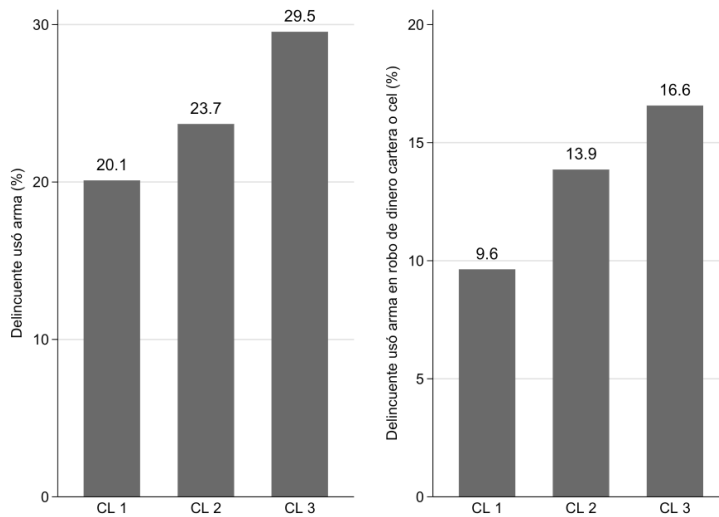
**Victimización indirecta y victimización repetida  
(en últimos doce meses) por clúster**



Fuente: ENAPRES, 2010-2016.  
Elaboración propia.

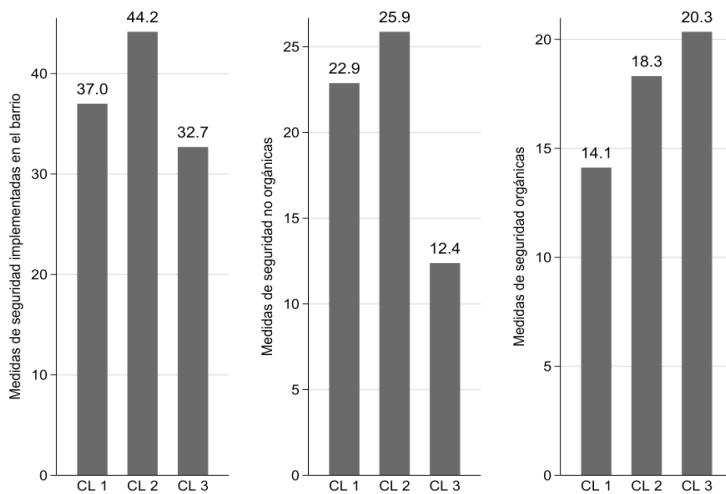
Figura 13.

**Victimización con arma (de fuego, cuchillo, palo, etc.) por clúster**



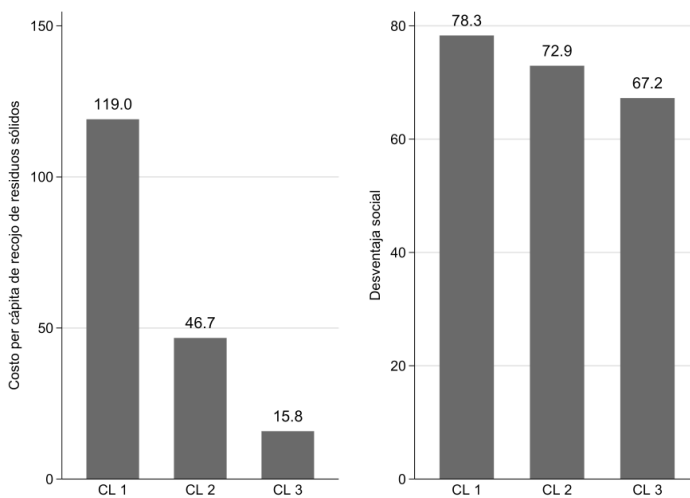
Fuente: ENAPRES, 2010-2016.  
Elaboración propia.

Figura 14  
Capital social por clúster



Fuente: ENAPRES, 2010-2016.  
Elaboración propia.

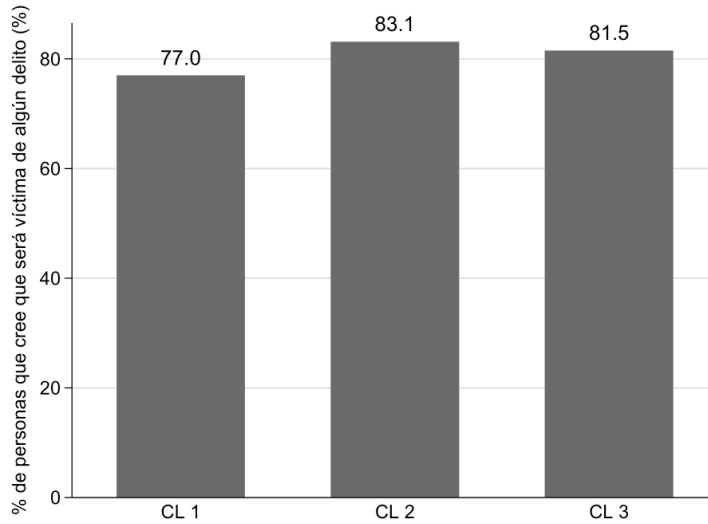
Figura 15.  
Desorganización social por clúster



Fuente: ENAPRES, 2010-2016.  
Elaboración propia.

Figura 16.

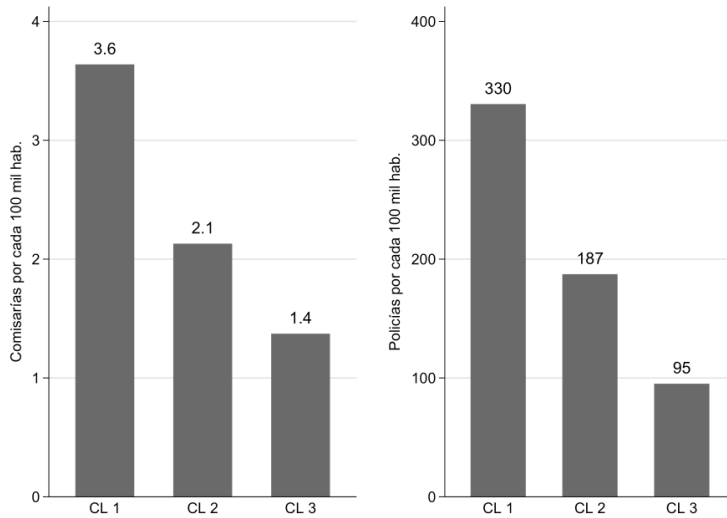
**Percepción de inseguridad por clúster**



Fuente: ENAPRES, 2010-2016.  
Elaboración propia.

Figura 17.

**Disponibilidad de comisarías y policías por clúster**



Fuente: ENAPRES, 2010-2016.  
Elaboración propia.

## Causas del crimen

### Metodología

La estructura de la data construida reposa en dos niveles de variables: individuales y distritales e individuales y clúster. Esta estructura anidada pone en evidencia una relación jerárquica entre personas y condiciones de ambiente, relación que se muestra adecuada para ser modelada como modelos jerárquicos lineales (multinivel). El supuesto detrás de esta elección es que el mayor o el menor grado de victimización en un distrito se debe tanto a las condiciones individuales de las personas que los habitan como a las condiciones sociales y económicas de los propios distritos. Pasar por alto la relación jerárquica entre individuos y distritos subvalúa las varianzas, al mismo tiempo que crea pruebas de significancia infladas con errores del tipo I (Williams, 2000) y estimadores insesgados y no eficientes (Goldstein, 2010).

En términos operativos, se parte de la ecuación:

$$victimización_{ij} = \beta_{0j} + \beta_1.X_{ij} + \varepsilon_{ij} \quad (1)$$

Nótese que la variable dependiente representa a la  $i$ -ésima persona víctima de algún hecho delictivo (cuando toma valor de 1) en el distrito  $j$ . A fin de captar la variabilidad a nivel distrital de los efectos sobre la victimización, el intercepto incluye el subíndice  $j$ , relación que podría reescribirse de la siguiente forma al permitir un intercepto común más una variable aleatoria:

$$\beta_{0j} = \beta_0 + \mu_{0j} \quad (2)$$

Al reemplazar (2) en (1), se observan las diferencias de la ecuación multinivel frente a una ecuación de mínimos cuadrados ordinarios. Frente a esta estimación, la multinivel permite descomponer la varianza de la victimización al nivel de las variables individuales y distritales.

$$victimización_{ij} = (\beta_0 + \mu_{0j}) + \beta_1.X_{ij} + \varepsilon_{ij} = \beta_0 + \beta_1.X_{ij} + \mu_{0j} + Z_j + \varepsilon_{ij} \quad (3)$$

En la ecuación anterior, son dos las variables aleatorias. Una es especificada para el nivel distrital (o para el clúster) ( $\mu_{0j}$ ) y la otra al nivel del individuo ( $\varepsilon_{ij}$ ). Los controles distritales son introducidos en el vector  $Z_j$ . Mientras que  $\mu_{0j}$  representa el vector de valores de interés, la lectura de las variables aleatorias permite realizar lecturas adicionales que indican la variabilidad de la variable dependiente al nivel individual y distrital (o clúster). Este modelo fue estimado con la data de 53,787 observaciones de individuos en 35 distritos de Lima; a la cual se le agregó información de cada uno de los distritos de los encuestados.

Existe un amplio consenso acerca de las fuentes de las causas del crimen. De un lado, hay ciertas características de las propias víctimas que elevan su probabilidad de convertirse en tales; pero estas no son los predictores más importantes. Incluso, las investigaciones seminales en desorganización social han estimado modelos que incluyen variables distritales. De otro lado, son ciertas condiciones y factores de riesgo en las comunidades las que aumentan la variabilidad del crimen en un barrio frente a otro (Sampson y Groves, 1989). Si se tiene en cuenta esto, a nivel individual se incluyó como controles al sexo y la edad del encuestado. También se consideró como control el mes en que se realizó la encuesta y la densidad poblacional de cada distrito.

La inclusión de variables a nivel contextual ha seguido nuestras tres teorías orientadoras. Se incluyó dos variables por cada teoría. La desorganización social fue evaluada mediante el índice de desventaja social<sup>1</sup>. En una estimación alternativa (no presentada) se incluyó el gasto municipal per cápita para el recojo de residuos sólidos (Registro Nacional de Municipalidades). Esta variable se aproxima a evaluar las condiciones del ornato público y su decaimiento sugiere una mayor presencia de puntos de riesgo. Sin embargo, no se incluyó por redundancia con el índice de desventaja social. El capital social se midió a través de dos variables. En primer lugar, se tomó el total de medidas de seguridad adoptadas en el barrio, dicho total varió entre cero y cinco y consideró las siguientes medidas: instalación de cámaras, contratación de vigilante particular, instalación de alarma comunitaria, instalación de tranquera y constitución de junta vecinal. En segundo lugar, se incluyó la confianza en la Policía. Finalmente, la teoría de las actividades rutinarias se midió mediante el número de comisarías y el número de policías por cada cien mil habitantes (Censo de Comisarías).

No se detectó multicolinealidad ( $VIF=2.13$ ) ni heterocedasticidad. El problema de endogeneidad fue resuelto mediante el uso de variables instrumentales. Al no contar con rutinas que incorporen instrumentos directamente en las estimaciones multinivel, se estimaron previamente modelos probit con las variables endógenas, sus instrumentos y los controles (sexo, edad y densidad). Luego se predijo el valor de la variable endógena, la cual fue incorporada en una segunda etapa en la estimación multinivel.

La relación entre victimización y desventaja social se da en doble sentido. Como no solo es posible que la desventaja social cree condiciones criminógenas, sino que la existencia de crimen deteriore y cree más desventajas sociales, se usó el número

1 Es un índice distrital estimado mediante análisis de componentes principales. Se empleó las siguientes variables (Censo Nacional de Población y Vivienda del 2007): porcentaje de mujeres con seguro de salud, prevalencia de pobreza, porcentaje de población mayor de edad, porcentaje de personas con castellano como lengua materna, porcentaje de personas residencialmente estables (viven más de 5 años en el distrito); porcentaje de personas en desempleo y porcentaje de jefas de hogar.



de licencias (de todo tipo) otorgadas por cada municipalidad como instrumento (Test de sobreidentificación de Sargan: p-value 0.995). La lógica de inclusión yace en que el número de licencias otorgadas se relaciona con la desventaja social pero no con la victimización. Los distritos que cuentan con la mayor cantidad de licencias emitidas generan una mayor sensación de orden y control, bien porque las nuevas edificaciones aumentan el flujo de gente o porque con estas se logra con una mayor cantidad de comercios. Paralelamente, se asume que la relación entre licencias municipales y crimen es nula, puesto que la evidencia previa sostiene que las licencias municipales no impactan sobre el desempeño de las empresas (microempresas) (Alcázar y Jaramillo, 2012) y este efecto genera más comercio en partes de la ciudad pero con un bajo incentivo para cometer delitos.

Es muy probable que la relación entre victimización y las medidas incluidas de capital social (medidas de seguridad tomadas en el barrio y confianza en la policía) también se influyen mutuamente. Para limpiar la endogeneidad se empleó como instrumento la variación distrital de la población (2010 y 2015) y la condición de jefe de hogar (dicotómica) del encuestado (Test de sobreidentificación de Sargan: p-value 0.818)

El crecimiento poblacional está relacionado con el aumento de la criminalidad, pero solo cuando este viene acompañado de una urbanización rápida y desordenada (Muggah, 2012). Sin esta condición, la relación se pierde. El sustento para el uso de este instrumento se debe en que Lima es una ciudad con un ritmo de crecimiento cada vez más lento, justamente porque su población ya es alta (y le cuesta más crecer) frente a otras ciudades. Además, en concordancia con lo que señala Ludeña (2002), no es el crecimiento urbano lo que ha determinado la condición de las ciudades en los últimos veinte años sino, principalmente, el sentido de precariedad y baja institucionalización como resultado de la débil consolidación de sectores sociales y sus intereses políticos y económicos. Como instrumento, la condición de ser jefe de hogar influye en el capital social (medidas de seguridad y confianza en la Policía), pero no directamente sobre la victimización, porque generalmente la jefatura del hogar recae en los hombres y esto es signo de una distribución desigual de los recursos del hogar que prioriza ciertas decisiones de gasto en este espacio por encima de otras.

Finalmente, se eligió al número de puestos de salud por distrito como instrumento para controlar la endogeneidad del número de comisarías y policías por cada cien mil habitantes con la victimización (Test de sobreidentificación de Sargan: p-value 0.187 y 0.223). Si bien es probable que se instalen más comisarías ahí donde la tasa delictiva es más alta, también se espera que se instalen donde haya o se necesite más presencia estatal. En estos últimos casos, la presencia de puestos de salud es una variable que se correlaciona con esa presencia estatal y no con la necesidad de instalar servicios estatales para lidiar con la delincuencia.

## Resultados

Se estimaron modelos multinivel con dos niveles para cada teoría. Los resultados se presentan en la Tabla 6 y en la Tabla 7. Ambas presentan la misma secuencia de cuatro estimaciones, con la diferencia que en la primera tabla se tomó al distrito como variable del segundo nivel (de agrupación de la data), mientras que en la segunda se tomó a los clústeres como variable de segundo nivel.

En general, los resultados muestran la importancia de las variables distritales sobre la influencia y variabilidad del crimen en Lima. Pero estas relaciones no se convierten en regla. Existen variables distritales sin influencia sobre la probabilidad de un individuo de ser víctima de algún hecho delictivo. Además, los efectos tienden a ser menos fuertes cuando se pasa de efectos de variables distritales a efectos de variables medidas en una agregación mayor como la del clúster.

El Modelo 1 valida la teoría de la desorganización social. En este, la desventaja social incrementa en forma significativa la incidencia de delitos. Debido a que la desventaja social es un indicador compuesto, este resultado sugiere que la combinación de ciertas condiciones sociales y económicas genera una asociación importante con la incidencia delictiva a nivel distrital. Si bien alguna de estas condiciones por sí sola también puede estar asociadas a una mayor criminalidad, nuestros resultados sugieren que su concurrencia genera una asociación importante. Además, lo relevante de la relación entre desorganización social y crimen se halla en que la primera refleja carencias que no necesariamente son consideradas como factores de riesgo por la literatura especializada. Por el contrario, la desventaja social se considera como un factor que debilita a otros, como la eficacia colectiva, hecho que a su vez influye en una mayor victimización.

A diferencia de la desorganización social, el capital social mostró un poder explicativo limitado al ser evaluado con el número de medidas de seguridad tomadas en el barrio y confianza en la Policía. En el Modelo 2 de la Tabla 6 se advierte que el efecto es positivo y significativo aunque bastante menor (marginalmente, igual a cero). Si se tiene en consideración que las medidas de seguridad incluyen cámaras de vigilancia, alarmas vecinales, juntas vecinales, tranqueras y vigilantes, estos resultados sugieren que tales elementos y dispositivos de seguridad se han convertido en respuestas rutinarias con menor impacto en la victimización. Probablemente se trata de medidas cuyo fin es proteger, pero cuya utilidad mayor podría hallarse en la creación de un falso efecto de tranquilidad.

En el otro caso, si se toma la confianza en la Policía como un factor que se aproxima a medir la percepción de la capacidad de control social desde el Estado, los resultados muestran que su contribución a la victimización es imperceptible. Al haberse descartado que la confianza en la Policía no explica más que marginalmente la victimización, surge la hipótesis adicional de que la confianza en la Policía se forma de expectativas poco focalizadas en la reducción de la victimización y que,

más bien, implica la materialización de expectativas distintas al objetivo de control social del delito y más cercanas a otras actuaciones policiales donde existe interacción con los ciudadanos, como la atención protocolizada en controles de tránsito (Murphy, Lorraine, y Bennet, 2013) y la satisfacción con la seguridad ciudadana y corrupción en funcionarios (Wu y Sun, 2009). Es oportuno que futuras investigaciones diferencien los distintos constructos detrás de la confianza en la Policía y los aislen de tal forma que la relación entre confianza en la Policía y victimización sea mejor comprendida.

Finalmente, la teoría de las actividades rutinarias falló en explicar la variabilidad de la victimización entre un distrito y otro. Se la operacionalizó como comisarías y policías por cada cien mil habitantes. En ambos casos, de acuerdo a lo esperado, en los distritos con más de estos recursos se redujo la probabilidad de victimización, siendo más fuerte la reducción por disponibilidad de comisarías que por cantidad de efectivos policiales. Sin embargo, también en ambos casos dichos efectos carecieron de significancia estadística. Dicho de otra manera, bajo la lógica de las actividades rutinarias, la disponibilidad de comisarías no actuaría como un guardián capaz, algo que sí es más fácil de lograr con mayores efectivos policiales.

Estos resultados sugieren que la presencia policial en Lima, sea a nivel de comisarías o policías, no es determinante como un factor que, entre un distrito u otro, genere variación en la probabilidad de victimización. Probablemente, otras características y cualidades del trabajo policial sean más importantes. Algunos estudios sostienen el efecto de más policías y menos crimen (Levitt, 2002), especialmente al nivel de ciudades (Marvell y Moody, 1996). Sin embargo, otros estudios señalan ciertas condiciones para que esta relación funcione. El efecto de más policías sobre menos crimen se daría ante eventos exógenos, como ataques terroristas (Di Tella, Galiani y Schargrodsky, 2010). Además, el efecto tampoco sería lineal. En línea con los resultados de Kleck y Barnes (2014), aumentar el número de policías afecta el nivel de crimen solo cuando se pasa de no tener policías a tener algo de efectivos, y no lo afecta cuando tales incrementos se dan sobre un stock de policías lejano al mínimo necesario.

Los resultados descritos líneas arriba fueron estimados con una estructura de individuos en distritos. La Tabla 7 cambia esa agrupación y estima individuos en clústeres; es decir, que mientras que en el primer caso la variabilidad de la victimización se estima en función de 35 distritos, en el segundo se hace con base en tres clústeres que agrupan a tal número de distritos. Los resultados son bastante similares a los de la Tabla 6, todos los signos, significancias y tamaños de los efectos se mantienen; la única diferencia importante es que el ratio de probabilidades de la desventaja social es menor. Esto quiere decir que en la medida en que la unidad geográfica se reduce, la variabilidad de la victimización amenora ante los cambios en la desventaja social. Este resultado sostendría la hipótesis de que en lugares con desigualdades menos pronunciadas el efecto de ciertas carencias sobre la victimización tiene menor importancia.

Tabla 6.

**Estimación de la victimización en Lima (odds ratio) (niveles: individuo/distrito)**

	(1)	(2)	(3)	(4)
Desorganización social				
. Desventaja social	12.53***			
Capital social				
. Medidas de seguridad tomadas en el barrio		0.000***		
. Confianza en la Policía		0.000***		
Actividades rutinarias				
. Comisarías por cada 100 mil hab.			0.882	
. Policías por cada 100 mil hab.				0.998
N	53,787	53,787	53,787	53,787
Controles	SI	SI	SI	SI
AIC	62,795	62,684	62,802	62,802
BIC	62,849	62,746	62,856	62,856

\*  $p < .1$  \*\*  $p < .05$  \*\*\*  $p < .01$

Tabla 7.

**Estimación de la victimización en Lima (odds ratio) (niveles: individuo/clúster)**

	(1)	(2)	(3)	(4)
Desorganización social				
. Desventaja social	8.191***			
Capital social				
. Medidas de seguridad tomadas en el barrio		0.000***		
. Confianza en la Policía		0.000***		
Actividades rutinarias				
. Comisarías por cada 100 mil hab.			0.910	
. Policías por cada 100 mil hab.				0.999
N	53,663	53,663	53,663	53,663
Controles	SI	SI	SI	SI
AIC	62,384	62,264	62,390	52,390
BIC	62,437	62,327	62,443	62,443

\*  $p < .1$  \*\*  $p < .05$  \*\*\*  $p < .01$

Los resultados anteriores consideran a los 35 distritos de Lima como un bloque, pero obvian que algunos de ellos son más parecidos entre sí, y que, por tanto, las causas del crimen pueden ser distintas. En razón de ello, en la Tabla 8 se presenta la misma estimación pero realizada en forma separada para cada uno de los tres clústeres previamente identificados (niveles, individuo y distrito). Cuatro aspectos resaltan de estas últimas estimaciones.

En primer lugar, los factores estadísticamente asociados a la victimización no son constantemente los mismos entre un clúster y otro. Este aspecto refuerza la lógica de distinguir los distritos de Lima en función de grupos de distritos que guarden algunas características similares. Ninguna variable fue consistentemente significativa entre los tres clústeres estudiados, lo que contribuye a la hipótesis de la heterogeneidad de las causas del delito en Lima y sugiere la necesidad de que las políticas tomen en cuenta las condiciones base de cada distrito (o grupo de distritos) para buscar resultados más efectivos.

En segundo lugar, solo el Clúster de Desprotección abierta, aquel que aglomera la mayor cantidad de distritos e indicadores de victimización y criminógenos más adversos, replica la significancia y signo de los resultados de la Tabla 7. En otras palabras, la desorganización tiene el mayor poder explicativo de la victimización, mientras que el efecto del capital social es positivo pero marginal y las actividades rutinarias no mostraron variables significativas. En el balance, estos resultados sugieren una situación delicada en el clúster de Desprotección abierta. Si consideramos únicamente los aspectos contextuales, los resultados sugieren que los factores de riesgo de victimización son mucho más fuertes que aquellos que buscan proteger a la población.

En tercer lugar, el único clúster donde la desorganización social no explica significativamente la victimización es en el de Protección latente. En este, los distritos gozan de mayores factores protectores, lo que probablemente signifique que la probabilidad de ocurrencia de un delito esté basada menos en las condiciones ambientales favorables del clúster y más en otras variables de teorías que enfatizan las características de los propios agresores y el desarrollo de su conducta delictiva y factores criminógenos (Andrew y Bona, 1994; Moffitt y Caspi, 1999), así como la vulnerabilidad y resiliencia de las víctimas (Walklate, 2011).

En cuarto lugar, algunas variables fueron significativas únicamente para algunos clústeres. Este resultado mostraría la particularidad de algunos clústeres y la fuerte variabilidad del efecto (positivo y negativo) que tienen algunas causas del crimen. Por ejemplo, las medidas de seguridad tomadas en el barrio solo explicaron significativamente la victimización en el clúster de Desprotección Abierta. Asimismo, la disponibilidad de comisarías solo tuvo un efecto negativo y significativo en el clúster de Protección Limitada.

En quinto lugar, las variables ligadas a las actividades rutinarias encontraron significancia únicamente en el clúster de Protección limitada. Una mayor disponibilidad de comisarías y de policías (en relación con el tamaño de la población) aumenta

significativamente la variabilidad de la victimización entre distritos. El efecto de la disponibilidad de comisarías fue incluso más fuerte que el de policías. Como se citó líneas arriba, la literatura que ha evaluado el efecto de más policías sobre el crimen no es concluyente, lo que sugiere que la falta de consistencia en tales resultados es más el reflejo de que el incentivo de más guardianes, visto desde la teoría de las actividades rutinarias, está muy influido por las condiciones locales del distrito o zona donde esta se aplique.

En resumen, ninguna teoría mostró consistencia a lo largo de cada clúster, ni cada clúster fue consistente a las mismas teorías. Estos resultados muestran heterogeneidad en las causas del delito en Lima y evidencian la condicionalidad del efecto de las características locales sobre la victimización, aunque también sugieren la existencia de otros factores explicativos que hacen alusión a la multicausalidad del crimen.

Tabla 8  
**Estimación de la victimización por clúster en Lima (odds ratio)**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	<b>Clúster 1: Protección latente</b>			<b>Clúster 2: Protección limitada</b>				<b>Clúster 3: Desprotección abierta</b>				
Desorganización social												
. Desventaja social	0.0001				1660383***				20.80***			
Capital social												
. Medidas de seguridad tomadas en el barrio		0.000				0.000				0.000***		
. Confianza en la Policía		0.547				28518.2				0.000***		
Actividades rutinarias												
. Comisarías (x 100 mil hab.)			0.064				1.687*				0.908	
. Policías (x 100 mil hab.)				1.083				1.008*				0.999
N	1011	1011	1011	1011	9473	9473	9473	9473	43179	43179	43179	43179
Controles	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI
AIC	909.8	908.4	909.8	909.8	10210.0	10193.5	10214.3	10214.3	51536.2	51449.6	51547.2	51547.2
BIC	929.5	933.0	929.5	929.5	10252.9	10243.6	10257.3	10257.3	51588.3	51510.3	51599.2	51599.2

\*  $p < .1$  \*\*  $p < .05$  \*\*\*  $p < .01$

## Conclusiones

Ante la ausencia de data distrital sobre victimización en Lima, este artículo representa un ejercicio de crearla y analizarla para abordar las causas del crimen, un asunto aún pendiente de respuesta. A partir de los siete años de la ENAPRES (2010-2016), se construyó un pool de datos (N=53,787) con la finalidad de responder dos preguntas: qué tan homogéneo es el crimen entre un distrito y otro en Lima (análisis de clúster) y analizar cuáles son las razones que hacen que un distrito en esta misma circunscripción tenga más victimización que otro (estimaciones multinivel). Estas preguntas estuvieron guiadas por las teorías que desde la Criminología han explicado la relación entre crimen y características del ambiente (desorganización social, actividades rutinarias y capital social).

La primera pregunta fue respondida empleando la data individual de 35 distritos de Lima. Nos quedamos con los distritos que, bajo el criterio de tamaño muestral y el coeficiente de disimilitud, aseguraron un error muestral y una variación aceptables. De este análisis se derivaron dos conclusiones importantes: que los índices de criminalidad revisados así como sus factores de riesgo presentan variaciones importantes entre un distrito y otro; y que la dispersión entre los distritos con mejores y peores indicadores en estos aspectos difieren en forma importante, incluso considerando los errores muestrales.

Ante la dificultad de analizar la distribución del delito en Lima, se construyeron clústeres de distritos que, basados en su similitud (respecto de indicadores de desorganización social, capital social y actividades rutinarias), pudieran reexpresar una agrupación más natural de los distritos limeños a partir de tales condiciones criminógenas. Este ejercicio produjo tres clústeres diferentes de distritos que denominamos Protección latente (N=3,551), Protección limitada (N=10,185) y Desprotección abierta (N=45,297). La conformación de estos grupos es consistente, tanto porque el primero presenta siempre los indicadores más positivos (y así sucesivamente), como porque esa posición la mantiene para casi todos los indicadores analizados.

Los distritos de Protección latente garantizan una protección no permanente. Si bien tienen tasas más bajas de victimización, victimización indirecta y victimización con armas más baja, tales indicadores son bajos en el contexto limeño de alta victimización y percepción de inseguridad. La protección es latente en estos casos porque existen recursos para afrontar la inseguridad aunque siempre sean insuficientes. Al igual que en este clúster, el clúster de Protección limitada ofrece ciertas condiciones para la seguridad; no obstante, el problema es la disponibilidad de comisarías y policías, así como la extensión de ciertas características (desorganización social) que crean condiciones para la aparición del delito. La protección, por tanto, se ofrece pero está limitada. Por último, el clúster de Desprotección abierta es el de desempeño más bajo. No solo cuenta con menos recursos, sino que la desorganización social es marcadamente más fuerte en este clúster. El estado en que la seguridad es



desplegada en estos casos, está más cercana de condiciones precarias que introducen permanencia en la protección.

La respuesta a la segunda pregunta planteada en este trabajo, las causas del delito en Lima, ofrece información para validar parcialmente las teorías testeadas.

La desorganización social se mostró como la más importante para explicar la victimización en Lima y reflejar la heterogeneidad del delito. En las estimaciones con todos los distritos fue la teoría con variables más importantes, pero dicha relación no siempre se mantuvo cuando las estimaciones se realizaron en forma independiente para cada uno de los clústeres en que se agruparon los distritos limeños. Nuestra hipótesis, de que la victimización está afectada por la desorganización social, no dejó de ser cierta, pero lo que los resultados sugieren es que la combinación de ciertas características sociales y económicas –resumidas en la desorganización social– crean condiciones de control y supervisión del crimen que no siempre aumentan la probabilidad de victimización ni su diferencia entre un distrito y otro. La desorganización social se activa como factor criminógeno ahí donde las condiciones sean adversas, como en los clúster de Protección limitada y Desprotección abierta.

El capital social tuvo un efecto no significativo o marginalmente nulo en todas las estimaciones, incluyendo las desagregadas por clúster. Pese a la evidencia previa que sí halla una relación positiva entre capital social y crimen, nuestro hallazgo se relaciona con la forma en que este ha sido medido en nuestra data. De un lado, las medidas de seguridad implementadas en el barrio se habrían convertido más en dispositivos de seguridad alineados con la necesidad de sentirse tranquilos que en la real protección que ofrecen frente actos delictivos; ese sentido instrumental y no funcional hace que su relación con el crimen pierda peso y sentido. En contraste, la evidencia señala que la asociación entre capital social y crimen se sostiene en un capital social con alta capacidad de organización y con capacidad de respuesta local (Bellair, 1997). Por otro lado, la confianza en la Policía también tuvo un efecto menor sobre el crimen, lo que sugiere que la confianza sea un concepto más asociado a otro tipo de actuaciones e interacciones entre policías y ciudadanos, como acciones de tránsito (Murphy, Lorraine, y Bennet, 2013) y satisfacción con la seguridad ciudadana o corrupción (Wu y Sun, 2009).

Finalmente, las actividades rutinarias señalan que la conjunción de agresores motivados, objetivos (víctimas) adecuadas y falta de guardianes incentiva la aparición del crimen. Al evaluar esta teoría a partir de la disponibilidad de comisarías y policías (por cada cien mil habitantes), ambas redujeron la variabilidad de la victimización pero siempre en forma no significativa. El único caso en que sí se halló significancia fue en las estimaciones por clúster donde, por el contrario, ambas variables aumentaron la ocurrencia de victimización. Este resultado se alinea con la no linealidad del efecto de contar con más policías. El efecto es más alto solo cuando se pasa de un escenario sin policía a otro con policía, y es pequeño o inexistente cuando el aumento va más allá de los niveles mínimos (Kleck y Barnes, 2014).

Este artículo no está libre de limitaciones. La opción tomada en este artículo fue considerar a la victimización como un bloque de distintas formas de hechos delictivos. Es probable que estas relaciones varíen en función del delito evaluado, enfatizando algunas características con la que se cometen ciertos hechos delictivos como la racionalidad del robo a hogares (Kelly, 2000) o la oportunidad del robo de vehículos (Nogueira, Adrensen y Fonseca, 2016). Además, el pool de datos construido para este estudio está sujeto a una variabilidad que bajo una muestra representativa distrital sería más fácil de controlar.

## Implicancias de política

La investigación pone énfasis en las distintas causas del crimen y en la heterogeneidad del mismo, así como en el funcionamiento (o no) de ciertas teorías explicativas de la victimización. Tales resultados nos permiten sugerir algunos puntos esenciales sobre cómo leerlos a la luz de las políticas existentes en materia de seguridad ciudadana.

En primer lugar, la asociación entre victimización y desorganización social, acentuada en los lugares donde esta última es más importante, implica la revisión de las políticas de seguridad y otras políticas sociales indirectamente relacionadas con condiciones criminógenas y de vulnerabilidad en las condiciones ambientales así como en las víctimas. Estrategias como Barrio Seguro, con enfoque multisectorial, han acertado al focalizarse en barrios con alta victimización y coincidentemente con mayores carencias sociales. Sin embargo, la desorganización social también alude a una baja capacidad del control social, característica no exclusiva de barrios de mayores carencias económicas. Pese a no concentrar altas tasas de victimización ni ser el objeto de programas particulares, algunos distritos se encuentran en un estado de protección limitada que puede ocasionar formas de victimización sostenidas en el tiempo. No es que se deban crear programas *ad hoc* para este tipo de distritos, sino que otros factores de la política pública deberían poner más atención a la formación de la victimización, como políticas de empleo juvenil y sobre todo otras que tengan influencia sobre el ambiente urbano (espacios públicos, iluminación, terrenos abandonados, etc.) y su efecto sobre oportunidades delictivas.

En segundo lugar, es importante promover desde las comisarías (Oficinas de Participación Ciudadana) la formación de un capital social más orgánico y sostenido en el tiempo. Las medidas de seguridad que normalmente adoptan los ciudadanos tienen un bajo impacto sobre su seguridad real, más allá de la propia percepción de seguridad que pueden generar. Lazos más orgánicos en la población permitirían generar medidas que, basadas en la cooperación y objetivo común de la seguridad, se alejen del paradigma de instalación de dispositivos de seguridad y se muevan hacia formas de cooperación vecinal y policial que mejoren los sistemas de identificación de zonas de riesgo, lugares peligrosos, población en riesgo y alertas sobre hechos delictivos. A

su vez, incorporar estos objetivos al trabajo de las comisarías podría mejorar el nivel local confianza ciudadana y le podría dar más contenido a esta a partir del objetivo de orden interno que tiene la Policía. Para ello, es necesario que se revalore en el propio Ministerio del Interior y Policía el trabajo de las Oficinas de Participación Ciudadana de las comisarías. Normalmente, son unidades con poco atractivo para los efectivos policiales y con limitaciones en sus modelos de gestión.

En tercer lugar, si bien es necesario aumentar la disponibilidad de policías y comisarías a fin de tener mayores puntos de contacto y denuncia hacia la ciudadanía, también es necesario reconocer que su aumento ha tenido un impacto nulo o incluso negativo sobre el crimen. A la luz de la evidencia sobre el efecto marginal o menor de incrementar el número de policías sobre el crimen, es importante revisar este enfoque. Diversos estudios señalan que la concentración de policías en puntos calientes o *hot spots* sí tiene efectos positivos en la reducción de la criminalidad, al igual que los que han estudiado la mejora de las estrategias de patrullaje. No solo es más policías, sino cómo se distribuye y utiliza tal recurso. Cada distrito debe actualizar sus mapas del delito y de riesgo dos veces al año. Sin embargo, mientras que los primeros se construyen con base en denuncias (sujetas a subreporte), los segundos se realizan con limitada participación de la población, pese a que estos dos instrumentos son vitales para determinar los lugares en los que se puede distribuir adecuadamente los recursos policiales con fines más efectivos.

En cuarto lugar, el Ministerio de Economía y Finanzas cuenta con una política de incentivos dirigida a gobiernos locales. Con cargo a cumplir algunos objetivos ligados a la instalación y funcionamiento de los Comités Locales de Seguridad Ciudadana, entre otros, los municipios reciben incentivos económicos. Sin embargo, en esta lógica subsisten tres problemas que pueden ser reorientados a la luz de estos resultados. Primero, los objetivos a cumplir son relativamente formales y simples de cumplir. Segundo, no hay la obligatoriedad de dirigir el incentivo económico a temas de seguridad. Tercero, las capacidades locales de elaborar planes de seguridad ciudadana basados en evidencia y bajo un enfoque multicausal son limitadas. Al revisar la evidencia disponible, es importante que se condicionen los incentivos señalados a programas y acciones nuevas o ya existentes que buscan reducir la desorganización social, que mejoran formas de capital social más sostenidas y que distribuyan en forma más adecuada los recursos policiales.

Esta lista de implicancias siempre resultará corta ante la multicausalidad de la violencia y las carencias institucionales. Lo importante es recordar que mientras más tarde se actúe, más difícil será controlar las condiciones que crean y perpetúan el crimen.

## Referencias bibliográficas

Agnew, R. (1992). Foundation for a general strain theory of crime and delinquency. *Criminology*, 30(1), 47-87.

Alcázar, L., y Jaramillo, M. (2012). *El impacto de la licencia municipal en el desempeño de las microempresas en el Cercado de Lima*. Lima: GRADE.

Andrews, D., y Bonta, J. (1994). *The psychology of criminal conduct* (5th ed.). New Jersey: Lexis Nexis, Anderson Publishing.

Barkan, S. (2012). *Criminology. A sociological understanding*. New Jersey: Prentice Hall.

Becker, G. (1968). Crime and punishment. An economic approach. *Journal of Political Economy*, 76(2), 169-217.

Bourdieu, P. (1986). *Distinction A Social Critique of the Judgement of Taste*. Londres, Routledge, 640 p.

Di Tella, R., Galiani, S., y Schargrodsy, E. (2010). Crime distribution and victim behavior during a crime wave. En R. Di Tella, S. Edwards, y E. Schargrodsy, *The Economics of Crime: Lessons for and from Latin America* (págs. 175-204). University of Chicago Press.

Goldstein, H. (2011). *Multilevel statistical models*. West Sussex: Wiley.

Kaplan, D. (2009). *Statistical Modeling. A fresh approach*. Project MOSAIC.

Kelly, M. (2000). Inequality and crime. *Review of economics and statistics*, 82(4), 530-539.

Kleck, G., y Barnes, J. (2014). Do More Police Lead to More Crime Deterrence? *Crime & Delinquency*, 60(5), 716-738.

Kubrin, C., y Weitzer, R. (2003). New directions in social disorganization theory. *Journal of Research in Crime and Delinquency*, 40(4), 374-402.

Kvålseth, T. (1995). Coefficients of Variation for Nominal and Ordinal Categorical Data. *Perceptual and Motor Skills*, 80(3), 843-847.

Le Blanc, M., y Fréchette, M. (1989). *Male criminal activity from childhood through youth*. (A. Blumstein, y D. Farrington, Edits.) New York: Springer-Verlag.

Levitt, S. (2002). Using electoral cycles in police hiring to estimate the effects of police on crime: Reply. *The American Economic Review*, 92(4), 1244-1250.

Ludeña, W. (2002). Lima: poder, centro y centralidad. Del centro nativo al centro neoliberal. *EURE*, 28(83).

Marvell, T., y Moody, C. (1996). Specification problems, police levels, and crime rates. *Criminology*, 4, 609-646.

Moffitt, T., y Caspi, A. (1999). Findings about partner violence from the Dunedin Multidisciplinary Health and Development Study. *National Institute of Justice. Research in Brief*, 1-12.

Mooi, E., y Sarstedt, M. (2010). *Cluster analysis*. Berlin, Heidelberg: Springer.

Muggah, R. (2012). *Researching the Urban Dilemma: Urbanization, poverty and violence*. Ottawa: International Development Research Centre.

Murphy, K., Lorraine, M., y Bennet, S. (2013). Promoting trust in police: findings from a randomised experimental field trial of procedural justice policing. *Policing and Society: An International Journal of Research and Policy*, 24(4), 405-424.

Nogueira, S., Adrensen, M., y Fonseca, L. (2016). Geography of crime in a Brazilian context: an application of social disorganization theory. *Urban Geography*, 1-23.

Sampson, R., y Groves, B. (1989). Community structure and crime: Testing social-disorganization theory. *The American Journal of Sociology*, 94(4), 774-802.

Santos, R., y Santos, R. (2015). An Ex Post Facto Evaluation of Tactical Police Response in Residential Theft from Vehicle Micro-time Hot Spots. *Journal of Quantitative Criminology*, 31(4), 679-698.

Walklate, S. (2011). Reframing criminal victimization: Finding a place for vulnerability and resilience. *Theoretical Criminology*, 15(2), 179-194.

Wikström, P.-O., y Loeber, R. (2000). Do disadvantaged neighborhoods cause well-adjusted children to become adolescent delinquents? A study of male juvenile serious offending, individual risk and protective factors, and neighborhood context. *Criminology*, 38(4), 1109-1142.

Williams, R. (2000). A note on robust variance estimation for cluster-correlated data. *Biometrics*, 56, 645-646.

Wu, Y., y Sun, I. (2009). Citizen trust in police. The case of China. *Police Quarterly*, 12(2), 170-191.

# Pasivos ambientales mineros, salud y agricultura: una primera aproximación espacial

José Carlos Orihuela<sup>2</sup>, César Huaroto<sup>3</sup> y Carlos A. Pérez<sup>4</sup>

## Resumen

En la literatura de desarrollo económico internacional abundan estudios sobre impactos nacionales y locales de las actividades extractivas en empleo e ingresos. No son frecuentes, en cambio, los estudios sobre externalidades negativas como la contaminación. Utilizando una base de datos administrativa sobre pasivos ambientales mineros (PAM) y encuestas de hogares georreferenciadas, exploramos cómo la cercanía a estos pasivos afecta las condiciones de salud y producción agropecuaria en hogares peruanos para el año 2016. Primero, obtenemos evidencia parcialmente robusta a correlación espacial de la relación entre cercanía relativa a un PAM e indicadores de salud: niños con menor altura y peso, según edad, y con mayor probabilidad de encontrarse desnutridos. Segundo, obtenemos evidencia que la cercanía a PAM se relaciona con un menor número de cabezas de ganado en hogares productores agropecuarios.

**Palabras Clave:** contaminación, pasivos ambientales mineros, salud, agricultura, geografía económica, maldición de los recursos naturales.

1 Esta investigación se desarrolló en el marco del Programa de Investigación Formativa y contó con la participación de los alumnos Rocío Arango y Gerald Cisneros.

2 PUCP. Contacto: orihuela.jc@pucp.pe

3 Universidad Católica de Chile.

4 PUCP.

## Introducción

Durante las últimas cuatro décadas se ha discutido y cuantificado los efectos económicos de la abundancia de recursos naturales, y se ha generado una extensa literatura sobre la llamada Maldición de los Recursos Naturales (MDR). Más recientemente, se observa un incipiente interés académico por estudiar los efectos económicos de la contaminación generada por las actividades extractivas (Aragón y Rud, 2015; Rau, T., Reyes, L. y Urzua, S.2015). Para el caso peruano, más allá de Herrera y Millones (2012), desconocemos de investigación económica cuantitativa sobre el tema. En parte, esta brecha en la literatura se debe a la poca disponibilidad de información sobre contaminación y la dificultad en poder establecer una relación causal desde un enfoque econométrico. Existe evidencia sobre efectos negativos de la contaminación minera en indicadores de salud (Rau et al. 2015, Menz y Seip 2004), así como efectos negativos en la producción agrícola (Aragón y Rud, 2015; Bebbington et al. 2013).

El objetivo de esta investigación es contribuir a la aún incipiente rama en cuestión, con un estudio exploratorio sobre las relaciones geográficas entre contaminación minera e indicadores de bienestar en Perú. Para lograr este objetivo, se utilizará un conjunto de herramientas que provee la geografía económica, y se combinará análisis espacial-geográfico (mediante Sistemas de Información Geográfica, GIS, por sus siglas en inglés) con el análisis estadístico econométrico. Con estas herramientas planteamos nuestra pregunta general de investigación: *¿cuál es la relación geográfica, en términos de cercanía y altitud, de un PAM con indicadores de salud y producción agrícola? Y a su vez una interrogante específica: ¿la cercanía a un PAM significa tener algún tipo de afección de salud y menor producción agrícola?*

La incorporación del componente espacial o geográfico es un aporte en sí mismo a la literatura en cuestión. La dimensión espacial está generalmente ausente de la investigación económica tradicional sobre la MDR (Orihuela, 2017), a pesar de que las características geográficas tienen una alta influencia en los resultados económicos de un entorno dado (Dell et al. 2014; Dell, 2010; Easterly y Levine, 2003). Para este tema particular, es relevante considerar la distancia entre un centro poblado y los pasivos ambientales mineros (PAM), así como la altitud, pues el encontrarse encima o por debajo de un pasivo tiene repercusiones diferentes. Si un PAM se encuentra encima de un centro poblado, la forma cómo afecta es del tipo *down-stream*, lo que facilita el esparcimiento de la contaminación vía aérea y del suelo; mientras que es más difícil que exista afectación en el caso que el centro poblado se encuentre por encima del PAM (Hendroychová y Kabra, 2016; Phillips, 2016). Con sus limitaciones, la altitud es una importante variable para incorporar al análisis.

El desarrollo empírico de la investigación es posible debido a la combinación de tres fuentes de información: (i) una base de datos administrativa provista por la Dirección General de Minería del Ministerio de Energía y Minas, en la cual se encuentra la ubicación, en términos de latitud y longitud, de 8,854 PAM; (ii) un conjunto de bases

de datos con información sobre indicadores de salud y producción agrícola que se encuentran georreferenciadas a nivel de centro poblado: la Encuesta Demográfica y de Salud Familiar 2016 y la Encuesta Nacional de Hogares 2016; y (iii) un Modelo Digital de Terreno (DEM, por sus siglas en inglés) sobre la altitud de cada 250 metros cuadrados del territorio peruano, que permite atribuir la altitud de cada PAM y centro poblado a fin de complejizar la relación geográfica propuesta.

En este estudio exploratorio obtenemos, por un lado, evidencia parcialmente robusta sobre la relación entre cercanía relativa a un PAM e indicadores de salud: niños con menor altura y peso según edad, y con mayor probabilidad de encontrarse desnutrido; por otro lado, obtenemos evidencia de que la cercanía a un PAM se relaciona con un menor número de cabezas de ganado en hogares productores agropecuarios. Estos resultados son discutidos en relación a los de otros estudios sobre el tema, los cuales no profundizan en el carácter espacial y podrían no ser robustos a correcciones aquí presentadas.

Nuestro trabajo cuenta con cinco secciones adicionales. La segunda sección expone el marco teórico, donde se discute la literatura sobre la maldición de los recursos referida a la contaminación generada por actividades extractivas. La tercera sección detalla los datos y metodología utilizados. En la cuarta sección se presenta un análisis descriptivo de los datos y en la quinta sección se expone los resultados. Finalmente, la sexta sección ofrece una breve discusión de los hallazgos, a modo de conclusión.

## Marco teórico

El desempeño económico de los países ricos en recursos naturales es un tópico de amplio interés y debate en las ciencias sociales (Cust y Poelhekke, 2015; Orihuela, 2017). Dentro de este debate, surge la paradoja de la MDR, la cual refiere que países ricos en recursos naturales tienen menor desarrollo debido a sus malos desempeños económicos, sociales y políticos. Un primer grupo de investigaciones de la literatura está compuesto por estudios econométricos *cross-country*, que han recibido críticas como el de adolecer de sesgos de variables no observables. Un segundo grupo de estudios busca superar las críticas mencionadas con investigaciones *within-country*, al analizar si dentro de un país existe MDR en aquellos lugares en donde se encuentran los recursos. Nuestro trabajo se inserta en este segundo grupo, y en particular, es parte de un conjunto de investigaciones exploratorias sobre los efectos de la contaminación de las actividades extractivas. A continuación, desarrollaremos sucintamente los grupos de literatura mencionados.

## Los estudios *cross-country*

Este primer grupo de estudios se ha desarrollado durante los últimos 30 años y de manera amplia. La literatura discute el rol de los recursos naturales en el desarrollo de los países, en particular si la abundancia de recursos naturales es una maldición en el



sentido de si ser rico en recursos naturales tiene un correlato negativo en indicadores económicos (posteriormente, la literatura incluirá resultados políticos y sociales). En la mayoría de trabajos se argumenta que existe una MDR, es decir que los países ricos en recursos suelen tener malos desempeños económicos.

Este grupo de estudios maneja un modelo teórico similar, buena parte referidos a la enfermedad holandesa (*Dutch disease*). Corden y Neary (1982) analiza una economía pequeña y abierta en la cual existe un *boom* en un sector primario, que trae como consecuencia el debilitamiento y reducción de la producción de sectores secundarios transables (como manufacturas) y el aumento de la producción de sectores no transables (servicios). El tipo de sectores económicos donde se dan los mecanismos favorece una inflación promovida por un exceso de demanda de productos no transables (servicios).

Se suele indicar a Auty (1993) como el pionero en acuñar la hipótesis de la “maldición” (Cust y Poelhekke, 2015; Orihuela, 2017). Posteriormente, Sachs y Warner (1995, 2001) continuaron y difundieron la idea. En 1995 Sachs y Warner encontraron que aquellos países ricos en recursos naturales (medido como tener un alto porcentaje de exportaciones primarias) tenían menores tasas de crecimiento. Posteriormente, Sachs y Warner (2001) hicieron un recuento de los estudios realizados previamente y defendieron la hipótesis de la MDR de las críticas sobre variables omitidas –geográficas o climáticas– que podrían explicar la MDR.

Los estudios *cross-country* han ofrecido un rico debate de ideas que ha intentado explicar la presencia de MDR en países ricos en recursos. El debate, más que conceptual, ha sido empírico y metodológico, pues los aportes diferenciales entre trabajos suelen ser el cambiar la variable de interés, la definición de ser rico en recursos y/o la inclusión de algún covariado ausente en especificaciones econométricas previas. Estos estudios plantean algunas limitaciones recogidas por Aragón, Pole y Land (2014): i) metodologías econométricas que no permiten solucionar aspectos como endogeneidad y errores de medición; ii) estudios que esconden patrones diferenciados de influencia de las actividades extractivas al interior de un país; y iii) investigaciones que no tienen claro el concepto de instituciones y determinar cuáles podrían ser las relevantes (a pesar de reconocer la importancia de las mismas).

### **Los estudios *within-country***

Recientemente, existen estudios sobre el impacto de ser rico en recursos naturales dentro de una región o áreas locales más desagregadas. Siguiendo a Aragón, Pole y Land (2014), estos estudios pueden dividirse según cuatro grandes grupos: los que analizan la dotación y especialización por los recursos, los que se interesan por los ingresos fiscales locales, los que se interesan por los shocks de demanda local, y los que se interesan por las externalidades negativas de las actividades.

El primer grupo de estudios investiga si el mecanismo de *Dutch disease* (“enferme-

dad holandesa”) funciona a nivel local, dado que las regiones pueden considerarse como economías pequeñas y abiertas. El *boom* de recursos es interpretado como un cambio de dotaciones, al cobrar los recursos un valor de gran magnitud. Las regiones se especializan en la explotación del recurso en *boom* y mellan la producción de otros sectores como manufacturas y agricultura, los cuales deben competir por los factores de producción trabajo, capital y tierra; mientras que sector de servicios gira en torno a la actividad en *boom* (Michaels, 2011).

El segundo grupo de estudios se concentra en los inusitados ingresos fiscales obtenidos por la extracción de recursos naturales: regalías, impuestos. Estos ingresos fiscales son rentas que los gobiernos nacionales obtienen fortuitamente, sin necesidad de una presión tributaria a los ciudadanos. Al haber mayores ingresos se puede brindar una mejor provisión de bienes públicos con lo que se mejoran las condiciones de vida de regiones que poseen los recursos. Sin embargo, estos ingresos pueden generar otro tipo de incentivos y una conducta buscadora de rentas que devienen en corrupción y mala asignación de recursos (Caselli y Michaels, 2013; Brollo et al., 2013).

El tercer grupo, muy ligado al primero, enfatiza que el sector en *boom* impacta sobre economías locales al comprar insumos locales como trabajadores y materiales intermedios (Aragón y Rud, 2013), ello genera un dinamismo económico de envergadura local a pesar de no ser muy fuerte a nivel macroeconómico (Phelps et al., 2015). En este grupo se interpretan las investigaciones a la luz de predicciones de teorías sobre mercados laborales locales y equilibrios espaciales (Moretti, 2011; Greenstone et al., 2010): salarios y empleo aumentan, se produce migración, se benefician sectores que producen insumos intermedios y el precio de viviendas aumenta debido a la migración.

El cuarto grupo analiza las externalidades negativas como la contaminación, que genera perjuicios en la salud y producción agrícola en personas y unidades agrícolas cercanas (Aragón y Rud, 2015; Rau et al., 2015; Bebbington, 2013). Este grupo de investigaciones será analizado en la siguiente sección.

Estos cuatro mecanismos expuestos resumen las investigaciones realizadas, las cuales frecuentemente enfatizan el carácter espacial de las investigaciones. El uso de Sistemas de Información geográfica (GIS) ha facilitado este tipo de análisis, ejemplo de esto son Kostadam y Tolonen (2016) y Aragón y Rud (2015, 2013).

Para Perú existen algunos trabajos dentro de los estudios *within country*<sup>1</sup>. El primer trabajo que utiliza los métodos resaltados para estos estudios es el de Zegarra et al. (2007), que investigan el modo en que la actividad minera impacta diferenciadamente en hogares urbanos y hogares agrícolas, y encuentra que hay un efecto positivo en hogares urbanos que no es muy claro en hogares agrícolas. Por otro lado, el estudio más conocido es el de Aragón y Rud (2013), quienes encuentran impactos positivos en indicadores de empleo e ingresos; particularmente, los autores estudian el efecto

1 Existen más estudios, que no han sido consideramos porque no han sido publicados como artículos de investigación sino que se encuentran como informes finales.

positivo que tuvo en una política de la empresa Yanacocha, de demanda de insumos locales. Escobal y Ticci (2015) utilizan censos de población y vivienda para el estudio y encuentran efectos positivos en empleo y reducción de pobreza en localidades andinas influenciadas por la minería. Del Pozo y Paucarmayta (2015) encuentran que minería y agricultura son actividades excluyentes en un mismo territorio, pues las unidades agropecuarias en el mismo ámbito geográfico de las mineras tienen menor producción agropecuaria. Loayza y Rigolini (2016) hallan que la pobreza se reduce en distritos mineros peruanos.

Se puede notar que estos trabajos son bastante empíricos y su discusión es más metodológica que teórica. La literatura no económica –entendida como aquella que no utiliza métodos cuantitativos o un marco centrado en el *Dutch disease* local– encuentra en esto una enorme debilidad, pues los resultados no pueden interpretarse y explicarse claramente, al ser una ‘caja negra’ (Orihuela, 2017; Thorp et. al., 2012). Incluso, de manera general, Angus Deaton –Premio Nobel de Economía 2015– critica esta forma de realizar investigación económica, pues para él es más importante entender el porqué de un fenómeno antes del qué (Deaton y Cartwright, 2016). Por otro lado, un aspecto no menor es que los estudios suelen tener conclusiones deterministas sobre la existencia o no existencia de MDR a nivel local.

## **Industrias extractivas y contaminación**

Las actividades extractivas, en particular la minería, genera contaminantes diversos que impactan sobre la productividad agrícola y la salud de las personas. Estas relaciones han sido aún poco estudiadas, y los estudios realizados son aún incipientes. A continuación, revisamos algunos trabajos acerca de esta temática.

Las actividades extractivas, debido a sus particulares procesos productivos, suelen contaminar el medio ambiente de zonas en donde se realiza la actividad (Aragón, Pole y Land, 2014). Algunas consecuencias de la contaminación son: la contaminación del agua utilizada para la agricultura (Bebbington, 2013), la contaminación del aire, que mella la productividad agrícola (Aragón y Rud, 2015), pues las siembras y el rendimiento de tierras a su vez son afectadas (Miller, 1988; Marshall et al., 1997). En general, la calidad de agua y de la tierra disminuyen (Menz y Seip, 2004).

Desde las ciencias ambientales, algunos estudios han utilizado enfoques experimentales y observacionales para analizar el crecimiento y rendimiento de cultivos ante la presencia de ciertos contaminantes como el dióxido de azufre (SO<sub>2</sub>), el ozono (O<sub>3</sub>) y las micropartículas (PM). Estos trabajos llegan a la conclusión que estos contaminantes son perjudiciales y llegan a disminuir los rendimientos entre un 37 y 42% según el tipo de cultivo (Emberson, 2001; Wahid et al., 1995). En particular, estos contaminantes afectan los cultivos de tres maneras: i) deterioran el aspecto físico de las plantas visiblemente, por lo que se dificulta su venta; ii) afectan directamente el crecimiento y el rendimiento; iii) alteran los procesos químicos y fisiológicos de los

cultivos y su entorno (Marshall et al., 1997). Los contaminantes mencionados suelen ser generados por los diversos procesos de extracción y producción minera (Aragón y Rud, 2015).

Por otro lado, diversos estudios encuentran que la contaminación afecta la salud de las personas expuestas, a las cuales provoca problemas de desnutrición y anemia (Goltz y Barnwal, 2013), disminución del desempeño educativo y posteriormente un bajo desenvolvimiento en el mercado laboral (Rau et al., 2015; Lavy et al., 2014) y menor productividad laboral (Hanna y Olivia, 2011; Graff y Zivin, 2011). De manera similar a lo que sucede en el caso de cultivos, los contaminantes alteran las capacidades fisiológicas del cuerpo humano e impiden su desempeño 'normal'. En particular, los niños menores de cinco años son vulnerables a la contaminación del medio ambiente, pues sus órganos se encuentran todavía en estado de formación, y especialmente su sistema respiratorio, el cual es vulnerable a contaminantes que suelen encontrarse cerca de los establecimientos mineros (OMS, 2017)

Los trabajos relacionados con las externalidades ambientales negativas de la actividad extractiva son más recientes y aún poco explorados. El marco teórico de estos y la justificación de los mecanismos se basan sobre estudios que examinan la contaminación en otros entornos, hecho que no los invalida, pues a pesar de provenir de diferentes fuentes emisoras de contaminación, analizan los mismos tipos de contaminantes. Goltz y Barnwal (2013) estudian la relación (para un conjunto de países 'en desarrollo') entre la cercanía a yacimientos mineros y los indicadores de salud, como anemia y desnutrición, en niños. Los resultados obtenidos muestran que la cercanía a estos centros tiene impactos negativos, pues aumenta la anemia y la desnutrición en los grupos mencionados, esto incluso a pesar del aumento de ingresos que genera la actividad productiva directamente. Se trataría de un *trade-off* entre beneficios económicos y efectos nocivos en salud, o como en el caso de La Oroya en Perú: un *trade-off* entre el derecho a trabajar y tener ingresos frente al derecho a la salud (Valencia, 2016). Rau et al. (2015) encuentran que los niños cuyas escuelas se encuentran en las cercanías de un botadero de residuos mineros tienen menores resultados educativos y sus ingresos de adultos en el largo plazo son menores. Aragón y Rud (2015), mediante el uso de información satelital sobre contaminación, encuentran que en Ghana la actividad minera trae consecuencias negativas para la productividad de pequeños agricultores familiares.

Otro tipo de literatura aborda el tema desde una perspectiva más cualitativa. Orihuela (2013) y Valencia (2016) analizan el caso de La Oroya con enfoques institucionales y de desarrollo de capacidades. Van Hoecke (2006), Ghose (2010) y Mucchloch (2010) analizan las dinámicas sociales de las mujeres en torno a la presencia de contaminantes. Otro tipo de literatura –no tan usual para el caso peruano– es aquel que valúa la contaminación ambiental minera, entendida como el costo que se incurriría para paliar la contaminación. Glave y Kuramoto (2002) señalan estudios de Evaluación Ambiental Territorial que realizaba el Ministerio; Herrera y Millones

(2012) se aproximan a los costos de la contaminación minera sobre los recursos hídricos para los años 2008 y 2009, y proponen, además, un esquema de multas y sanciones ambientales.

## La maldición de los recursos naturales

En los diversos tipos de literatura revisados, las conclusiones obtenidas por los autores suelen ser determinísticas: existe (o no) MDR debido a un resultado en una variable determinada. Siguiendo a Orihuela (2017), creemos que la MDR es un concepto de tipo ‘sombrija’ que sirve para llamar a un conjunto diverso e interrelacionado de enfermedades que podría conllevar el desarrollo basado en recursos naturales. La MDR no debe ser interpretada de manera universal, ni determinista. La multidimensionalidad del desarrollo plantea situaciones en las que puede haber ‘maldición’ en un aspecto y ‘bendición’ en otro; por ejemplo, el último *boom* de precios de minerales trajo una ‘bendición’ macroeconómica, pero una ‘maldición’ social al agravarse los conflictos sociales. Por otro lado, una maldición podría volverse bendición, y viceversa.

Consideramos que el concepto de MDR tiene una taxonomía particular que conlleva retos de gobernanza en aspectos macroeconómicos, de sectores y economías locales, de sostenibilidad del crecimiento económico; de manejo del control democrático y político; de justicia ambiental y de paz y manejo de conflictos. Cada uno de estos desafíos de gobernanza está asociado con distintas ‘enfermedades del desarrollo’ y con mecanismos peculiares (Orihuela, 2017).

Particularmente, este trabajo se enfoca en los retos de gobernanza de justicia ambiental y de paz. La ‘bendición’ en empleo e ingresos que puede generar la actividad extractiva puede conllevar una ‘maldición’ al emitirse externalidades negativas (contaminación) que afectan el desempeño de otras actividades económicas (agricultura) y la salud de las personas. Esta convivencia entre maldición y bendición genera el *trade-off* discutido anteriormente (Valencia, 2016), y plantea un reto fundamental de gobernanza para un país con una economía dependiente de estas actividades, con diversos conflictos sociales generados en torno a esta actividad y a la vez con 8,854 PAM a lo largo de su territorio.

## Fuentes de información y metodología

El objetivo de este trabajo es proveer un análisis exploratorio sobre las relaciones geográficas entre contaminación minera e indicadores de bienestar en el Perú. Al ser un tema nuevo en la agenda de investigación del medio y por características propias de la información disponible, el presente trabajo tiene un carácter exploratorio y brinda conjeturas sobre el modo en que indicadores de salud e indicadores de producción agrícola se encuentran relacionados al hallarse ‘cerca’ de un PAM. Investigaciones detalladas en la sección anterior utilizan la cercanía –a un centro minero o a una fuente de contaminación minera– como una distancia euclidiana (Aragón y

Rud, 2015; Kostadam y Tolonen, 2016; Rau et al., 2015; entre otros), pero obvian que la altitud juega un rol importante, sobre todo en contextos geográficos como la sierra peruana, en donde la ausencia de llanuras invalidaría o cuestionaría el uso de distancias euclidianas.

La hipótesis principal de esta investigación es que la cercanía a una PAM tiene una relación negativa en indicadores de salud y producción agrícola, tal como ha encontrado la literatura reseñada en la sección 2.3. Además, la relación se intensifica al considerar un componente geográfico importante como la altitud (Phillips 2016, Emel et al 2014); particularmente, cuando la fuente de contaminación se encuentra por encima de las personas y unidades agrícolas, la relación tipo *downstream* favorece la expansión de la contaminación.

Se corroborará la hipótesis combinando técnicas de análisis espacial (típicamente utilizados por geógrafos) y análisis estadístico-econométrico, aprovechando tres fuentes de información debidamente georreferenciadas. En las siguientes sub-secciones describiremos estas fuentes y explicitaremos la metodología de análisis.

## **Fuentes de información**

Esta investigación utilizará tres fuentes de información, las cuales tienen en común el encontrarse debidamente georreferenciadas en términos de longitud y latitud, así como de altitud para las dos últimas. La primera fuente será proyectada y junto a un DEM de altitud se podrá extraer las características de altitud requeridas para el análisis geográfico.

La primera fuente de información es el registro de Pasivos Ambientales Mineros (PAM) provisto por la Dirección General de Minería del Ministerio de Energía y Minas (MINEM). Esta información se ha publicado en la página web del MINEM, pero no en formato de base de datos. Esta base de datos es una especie de censo de PAM que se encuentran tipificada y georreferenciada. El no contar con información sobre la fecha en que se originó el pasivo o una medida sobre el nivel de contaminación o riesgo de cada pasivo imposibilita realizar un análisis más fino. No obstante, se selecciona ciertos PAM, pues existe un grupo que incluye infraestructura minera no contaminante como oficinas: i) del tipo labor minera seleccionamos los tajos; ii) del tipo productos químicos seleccionamos todos: aceites, combustibles y cianuros; y, iii) del tipo residuos mineros seleccionamos todos: desmontes mineros, relaves mineros, lodos de neutralización, material de desbroce, pilas de lixiviación, residuos de carbón y suelo orgánico.

La segunda fuente de información es la Encuesta Nacional de Hogares 2016 (ENAH), la cual provee un módulo detallado sobre Producción Agropecuaria de Hogares cuyos miembros se dediquen a la agricultura como actividad principal, que son la mayoría de hogares rurales. Esta encuesta cuenta con una georreferenciación bastante precisa a nivel de centro poblado (para zonas rurales) y a nivel de manzana (para zonas urbanas).

La tercera fuente de información es la Encuesta Demográfica y de Salud Familiar 2016 (ENDES) que provee indicadores de salud objetivos (talla, peso, hemoglobina) para niños menores de cinco años, y que posee la georreferenciación a nivel de centro poblado<sup>2</sup>.

Una fuente adicional que se usa para obtener la altitud a la que se encuentran los PAM es un DEM de altitud (el Shuttle Radar Topography Mission -SRTM) desarrollado por la NASA y el Ministerio de Economía de Japón y se encuentra en formato de ráster<sup>3</sup>. Este, proyectado en un GIS es una imagen en la cual cada píxel de resolución representa una información distinta sobre la altitud. El píxel cuenta con una resolución de 250 metros cuadrados, y al superponer la capa que provee la ubicación de los PAM se puede obtener la altitud de cada uno de estos. Con este DEM se calcula también la pendiente de los centros poblados y los PAM, pues indica lo rápido que se puede esparcir una fuente contaminante. Debe considerarse que existen DEM con mayor precisión, como de 30 y 90 metros cuadrados; sin embargo, estos requieren un tratamiento especial para llenar 'huecos' de información en ciertas nubes, el SRTM a 250 metros facilita el análisis.

## Metodología

Dados los objetivos del estudio y el carácter exploratorio de la investigación, la metodología a seguir constará de un análisis de regresión de corte transversal cuya interpretación causal es discutible, debido a falta de información como una medida objetiva sobre contaminación en los PAM, desde la fecha que existen, problemas de endogeneidad, entre otros. A pesar de estas limitaciones, se plantean diversos modelos basados en estudios especializados acerca de los temas que permiten incluir las variables relevantes que controlen las características que influyen en los indicadores seleccionados.

Los modelos presentados tienen como interés principal determinar cómo afecta la 'cercanía' de un PAM en indicadores de agricultura y de salud de un centro poblado. Estudios previos definen esta cercanía como la distancia euclidiana, la cual, dado el contexto geográfico de la sierra debe reconsiderarse si se toma en cuenta la altitud relativa y la pendiente del suelo donde se ubican los PAM. La altura relativa tiene importancia, pues si la fuente contaminante se encuentra más elevada que un centro poblado, podría producir filtraciones que podrían trasladarse por tierra, agua y aire con mayor facilidad (Hendrychova y Kabrna, 2016; Phillips, 2016; Dufflo y Pande, 2005). Dicho esto, el modelo general es el siguiente:

$$y_{ic} = \alpha_0 + \beta(\text{cerca}_c \times \text{pam\_alto}_c) + \beta_0 \text{cerca}_c + \beta_1 \text{pam\_alto}_c + X'_{ic} \theta + \epsilon_{it}$$

<sup>2</sup> Si bien la ENDES provee otros indicadores de salud que suelen utilizarse en otros estudios, estos indicadores son subjetivos o de autorreporte, y existe una gran posibilidad de que contengan errores de medición; estos autorreportes se refieren a enfermedades diarreicas y respiratorias. Las medidas de hemoglobina y antropométricas permiten un acercamiento más objetivo.

<sup>3</sup> El formato ráster es una imagen que puede entenderse como la combinación de píxeles. Cada píxel tiene un valor que refleja el atributo de la localización del píxel. Para nuestro caso, cada píxel representa la altitud y cuenta con una resolución de 250 metros cuadrados.



Donde 'y' es la variable dependiente de interés para el individuo 'i' ubicado en el centro poblado 'c'. El coeficiente de interés es  $\beta$ , de la interacción entre las variables cerca y pam\_alto, *dummies* que toman el valor de 1 en caso el centro poblado 'c' donde vive el individuo 'i' se encuentre 'cerca' de un PAM y a una altura menor que el del PAM 'cercano'. X representa un conjunto de características de cada individuo,  $\epsilon$  representa los errores del modelo. No existe un criterio para definir qué significa estar 'cerca' a un PAM, por lo que se proponen tres distancias (5, 10 y 15 kilómetros) y se espera que el efecto sea mayor a mayor cercanía. Debe resaltarse que el coeficiente de interés, al ser una interacción de variables 'dummies' debe interpretarse como la diferencia o comparación entre individuos 'cercanos' y ubicados por debajo de un PAM respecto a individuos 'cercanos' pero ubicados por encima de un PAM.

Existe multicausalidad en los desempeños de salud de niños menores de 5 años, al igual que en el valor de producción de pequeños productores. Por ese motivo, es importante elegir un conjunto de variables de control que la literatura especializada indique como relevantes a fin de tener estimadores insesgados; es decir, considerar otras variables que pueden ser explicativas sobre la salud y la producción agropecuaria. Para lograr este objetivo seguimos las especificaciones y variables de control de dos interesantes estudios, en los cuales se discute la importancia de incluir las variables mencionadas.

Para el caso de los indicadores de salud, seguimos a León et al. (2016), quienes estudian el efecto de la violencia doméstica en la salud de niños menores de 5 años. Los autores utilizan la base de datos ENDES y las mismas variables dependientes, por lo que su estudio representa un buen 'modelo' a seguir. Las variables dependientes son 'objetivas', en el sentido que se basan sobre mediciones antropométricas y no autorreportes o percepciones sobre salud, y son: i) índices estandarizados de mediciones alternativas de peso y talla, de acuerdo con criterios de la Organización Mundial de la Salud (OMS); ii) mediciones de hemoglobina y su derivación en condición de anemia. Por otro lado, las variables de control escogidas son características relevantes que también podrían explicar estos indicadores de salud: i) características del niño: edad en meses y sexo; ii) características de la madre: edad, escolaridad, condición laboral y condición conyugal; iii) características del hogar: resumidas en un índice de riqueza.

Por otro lado, para el caso de los indicadores agropecuarios, seguimos a Aragón y Rud (2015), quienes estudian el efecto de la contaminación aérea de la minería sobre la productividad agropecuaria en Ghana. El indicador agropecuario escogido es el valor de producción agropecuaria y la mejor forma de explicarlo es bajo un enfoque de función de producción, considerando los factores de producción utilizados: tierra, mano de obra (familiar y contratada) e insumos intermedios. Otras características relevantes son las referentes al productor agropecuario (sexo, edad, estado civil, nivel educativo, distrito de nacimiento) y a la unidad agropecuaria (superficie titulada, superficie bajo riego).

Aparte del modelo propuesto inicialmente, se realizarán modificaciones que incluyen ciertos efectos fijos para probar la robustez de los resultados. En particular, debe considerarse que hacemos uso de datos georreferenciados, por lo que es esperable



encontrar correlación espacial (Dell, 2010; Dell et al., 2014) que debe tenerse en cuenta al momento de discutir los resultados. Tradicionalmente, los estudios no han prestado mucha atención a este carácter, a pesar de que en presencia de correlación espacial, los estimadores obtenidos podrían ser insesgados e inconsistentes (Anselin, 2010; Lesage y Pace, 2014); ello debe tenerse en cuenta.

Presentamos dos especificaciones que suelen realizarse para corregir la correlación espacial, discutidos en Hsiang (2016) quien hace un recuento sobre los estudios econométricos sobre clima, contexto con alta correlación espacial. Una primera manera es agrupar los errores estándares a un nivel desagregado como el distrital (*clusterized standard errors*), y otra alternativa es presentar errores estándares siguiendo a Conley (1999). Siguiendo a Hsiang (2016), preferimos la primera alternativa (*clusterized standard errors*) debido a que con la segunda podría haber dificultades en estimar una matriz de varianza-covarianza bien definida (Hsiang 2016).

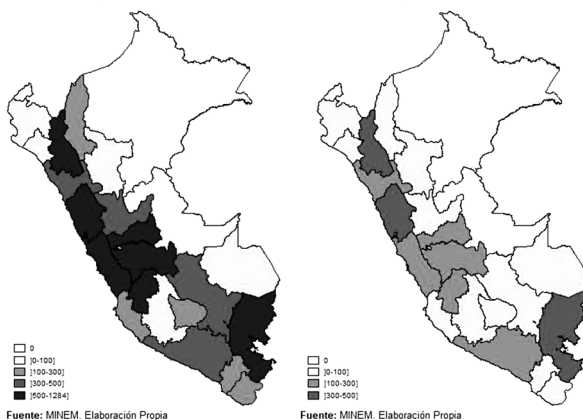
## Análisis descriptivo

### Ubicación de PAM

Como se ha mencionado, en el Perú existen 8854 PAM a lo largo de 21 departamentos. Los departamentos en los cuales no existen son Tumbes, Loreto y Ucayali. Estos pasivos incluyen infraestructura que no puede considerarse como contaminante, por lo que se hace una selección de un subgrupo de estos (ver sección anterior) y se obtienen 3021 PAM considerados como potencialmente contaminantes y dañinos. Los departamentos con mayor número de pasivos seleccionados son Cajamarca, Áncash y Puno, con más de 300 cada uno.

Figura 1

### Mapas de Pasivos ambientales mineros (PAM) por departamento

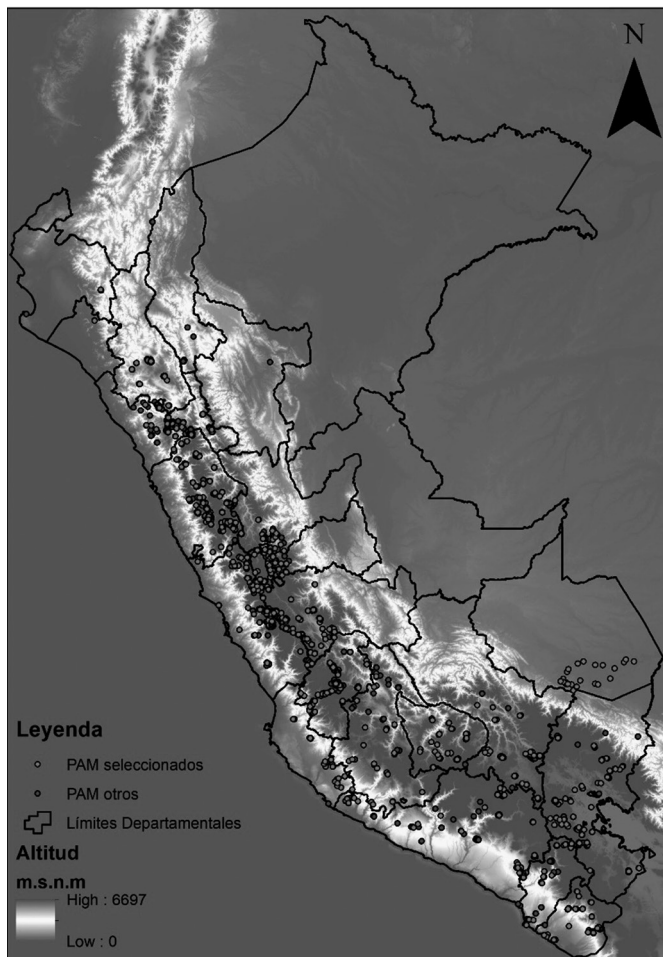


Mapa 1. PAM totales.

Mapa 2 PAM contaminantes.

Los mapas 1 y 2 proveen información interesante, aunque no muestran la distribución de los pasivos en el espacio de manera más detallada. En el Mapa 3 presentamos la ubicación de cada PAM en el espacio, con una capa previa que brinda información sobre la altura de estos. En este mapa, podemos rescatar que la mayoría de los PAM se encuentran en zonas altas, es decir la sierra peruana. Este resultado es esperable, debido a que los minerales, por motivos geológicos se encuentran en estas zonas (Zegarra et al., 2007). Del Mapa 3 se desprende una observación interesante que puede ser obviada en mapas como los anteriores mostrados: no todo el departamento se encuentra afectado por los PAM sino ciertos espacios particulares dentro de estos.

Figura 2  
**Mapa de ubicación de PAM y su altitud**



Fuente: MINEM y SRTM. Elaboración propia.

## Indicadores de salud

La fuente de información utilizada para analizar los indicadores de salud es la ENDES 2016, que entre otros indicadores provee indicadores objetivos sobre la salud de niños menores de 5 años. En la Tabla 1 presentamos los estadísticos descriptivos de las variables utilizadas. Se puede notar que según distancia, los niños no tienen características totalmente diferentes.

Tabla 1.  
Estadísticos descriptivos. ENDES 2016

Variables	Distancia 5 km			Distancia 10 km			Distancia 15 km			Distancia > 15 km		
	N	Media	Sd	N	Media	Sd	N	Media	Sd	N	Media	Sd
<b>Indicadores de resultado</b>												
Hemoglobina	1724	11.06	1.20	3523	11.15	1.20	4811	11.18	1.21	7567	11.38	1.16
Anemia	1724	43.91	49.64	3523	40.70	49.14	4811	40.43	49.08	7567	33.53	47.21
Altura según edad	1846	-1.00	1.00	3770	-0.96	1.00	5149	-0.99	1.00	8083	-0.92	1.05
Peso según edad	1846	-0.31	0.98	3770	-0.22	1.01	5149	-0.23	1.01	8083	-0.10	1.08
Peso según altura	1846	0.36	0.96	3770	0.45	0.98	5149	0.46	0.98	8083	0.59	1.02
Desnutrición	1846	14.68	35.40	3770	13.66	34.35	5149	14.31	35.02	8083	14.82	35.53
Bajo Peso	1846	3.68	18.84	3770	3.18	17.56	5149	3.34	17.97	8083	2.98	17.01
<b>Características niño</b>												
Edad en meses	1865	29.91	16.53	3819	30.14	16.58	5223	30.14	16.66	8159	30.53	16.64
Sexo (=1 Mujer)	1865	0.47	0.50	3819	0.48	0.50	5223	0.48	0.50	8159	0.48	0.50
<b>Características madre y hogar</b>												
Años educación	1838	3.92	1.47	3751	3.94	1.47	5124	3.97	1.47	7980	4.01	1.49
Índice riqueza	1865	0.14	0.78	3819	0.18	0.88	5223	0.12	0.90	8159	0.03	0.97
Edad	1865	29.75	6.69	3819	30.06	6.83	5223	30.09	6.85	8159	30.06	7.05
Jefe de hogar	1865	0.13	0.34	3819	0.12	0.33	5223	0.12	0.32	8159	0.11	0.31
Tiene pareja	1865	0.85	0.36	3819	0.86	0.35	5223	0.86	0.35	8159	0.87	0.34
Trabaja	1865	0.53	0.50	3819	0.56	0.50	5223	0.57	0.50	8159	0.56	0.50

## Indicadores de producción agropecuaria

La fuente de información utilizada para analizar los indicadores de producción agrícola es la ENAHO 2016, en particular el módulo de Productor Agropecuario, que entre otros provee indicadores de producción agropecuaria y factores de producción utilizados. En la Tabla 2 presentamos los estadísticos descriptivos de las variables utilizadas. Se puede notar que según distancia, los hogares productores no tienen características totalmente diferentes.

Tabla 2  
Estadísticos descriptivos. ENAHO 2016

Variables	Distancia 5 km			Distancia 10 km			Distancia 15 km			Distancia > 15 km		
	N	Media	Sd	N	Media	Sd	N	Media	Sd	N	Media	Sd
<b>Variables de resultado</b>												
Prod. agrícola (Ln)	685	7.18	1.61	1850	7.19	1.62	3035	7.23	1.52	6497	7.53	1.62
Prod. pecuaria (Ln)	742	6.10	2.48	1925	6.08	2.54	3113	6.05	2.57	6608	6.09	2.42
# Ganado	742	2.27	1.62	1925	2.22	1.65	3113	2.21	1.65	6608	2.32	1.58
Prod. agropecuaria (Ln)	854	7.63	1.88	2172	7.76	1.75	3467	7.80	1.73	7406	7.96	1.75
<b>Factores de producción</b>												
Superficie agropecuaria (Ln)	854	-0.61	1.40	2172	-0.46	1.38	3467	-0.39	1.33	7406	-0.12	1.40
Gasto en L agrícola (Ln)	686	0.65	3.90	1855	0.82	3.89	3043	0.69	3.83	6508	1.01	4.08
L familiar	854	-1.89	1.01	2172	-1.79	1.09	3467	-1.74	1.14	7406	-1.81	1.07
Gasto en insumos agrícolas	686	6.11	1.63	1855	6.10	1.60	3043	6.12	1.54	6508	6.24	1.77
Gasto en L pecuario (Ln)	706	-1.80	2.04	1849	-1.92	1.74	2990	-1.99	1.56	6416	-2.09	1.28
Gasto en insumos pecuarios	706	5.65	1.57	1849	5.66	1.58	2990	5.67	1.57	6416	5.70	1.52
Gasto en L agropecuario	854	0.38	3.86	2172	0.55	3.85	3467	0.47	3.79	7406	0.70	4.01
Gasto en insumos agropecuarios	854	6.34	2.03	2172	6.46	1.86	3467	6.51	1.80	7406	6.68	1.81
<b>Características hogar - Unidad agropecuaria</b>												
Sexo (Mujer)	854	0.25	0.43	2172	0.24	0.43	3467	0.24	0.43	7406	0.19	0.39
Edad	854	55.06	14.59	2172	54.67	15.19	3467	54.50	15.38	7406	53.83	15.55
Edad 2	854	3244.82	1624.38	2172	3219.17	1702.47	3467	3206.67	1722.60	7406	3199.89	1736.95
Nació en mismo distrito	854	0.88	0.32	2172	0.91	0.29	3467	0.91	0.28	7406	0.88	0.33
% Tierra titulada	699	16.23	35.68	1884	14.74	33.97	3084	13.32	32.47	6587	14.85	34.53
% Tierra bajo riego	699	46.02	46.51	1884	50.44	46.31	3084	49.25	46.19	6587	42.23	46.40

Variables	Distancia 5 km			Distancia 10 km			Distancia 15 km			Distancia > 15 km		
	N	Media	Sd	N	Media	Sd	N	Media	Sd	N	Media	Sd
<b>Estado civil</b>												
Conviviente/Casado	854	0.71		2172	0.71		3467	0.71		7406	0.75	
Viuado	854	0.15		2172	0.15		3467	0.15		7406	0.13	
Separado/Divorciado	854	0.09		2172	0.09		3467	0.09		7406	0.08	
Soltero	854	0.05		2172	0.05		3467	0.05		7406	0.04	
<b>Nivel educativo</b>												
Sin nivel	854	0.21		2172	0.20		3467	0.19		7406	0.15	
Inicial	854	0.12		2172	0.15		3467	0.15		7406	0.16	
Primaria completa	854	0.28		2172	0.30		3467	0.32		7406	0.37	
Secundaria completa	854	0.39		2172	0.36		3467	0.34		7406	0.32	
Superior completa	854	0.15		2172	0.12		3467	0.11		7406	0.09	
Postgrado	854	0.01		2172	0.01		3467	0.01		7406	0.00	

## Resultados

Los resultados obtenidos indican que la cercanía a un PAM afecta negativamente algunos indicadores de salud y de producción agropecuaria. Para la facilidad de lectura, presentamos el coeficiente de interés para tres bloques, el cual resulta de la interacción entre la cercanía y la altitud relativa, según la distancia definida. El coeficiente mostrado compara la diferencia entre quienes están ‘cerca’ y debajo del PAM con aquellos que están ‘cerca’ y encima de un PAM.

### Indicadores de salud

Nuestra primera aproximación a los efectos de los PAM sobre el bienestar de personas cercanas es considerando indicadores de salud de niños menores a 5 años. En particular, elegimos siete indicadores: nivel de hemoglobina ajustado por altura, probabilidad de tener anemia (medida según el nivel de hemoglobina), los índices estandarizados de la Organización Mundial de la Salud (OMS) para talla según edad (HA), peso según edad (WA) y peso según talla (WH), y la probabilidad de encontrarse desnutrido y bajo de peso (según los estándares de la OMS a partir de los índices de HA y WA).

El modelo principal considera las variables de control propuestas por León et al., (2016), efectos fijos a nivel de provincia y controles por el grado de pendiente del PAM ‘cercano’ y el centro poblado. En la Tabla 3 se muestra los resultados para las definiciones de ‘cercanía’ de 5, 10 y 15 kilómetros de distancia entre centro poblado y PAM. El coeficiente de interés es aquel sobre la interacción entre la definición de cercanía y la condición de que el PAM se encuentre a mayor altura que el centro poblado. En adelante llamaremos a la interacción ‘cercanía relativa’. Restringimos la muestra a aquellos centros poblados ubicados a menos de 100 kilómetros de un PAM y no consideramos Lima Metropolitana, por ser un caso particular.

Encontramos que el hecho que hallarse relativamente cerca de un PAM tiene efectos negativos en la altura y peso según edad de los niños menores de 5 años. Los niños que se encuentran a 5 kilómetros de un PAM tienen una altura para la edad menor en -0.121 desviaciones estándar y -0.110 desviaciones estándar en peso para la edad. Considerando una cercanía de 10 kilómetros, se obtiene -0.136 para altura según edad y -0.098 para peso según edad; y además que la probabilidad de encontrarse desnutrido aumenta en 5%. No se encuentran resultados significativos al considerar una distancia de 15 kilómetros. Los demás indicadores (hemoglobina, anemia, entre otros) no tienen significancia estadística; sin embargo, debe rescatarse que el coeficiente tiene un signo teóricamente esperable: positivo en la probabilidad de anemia, desnutrición y bajo peso, y negativo en hemoglobina y peso según altura.

Tabla 3.  
**Resultados principales de cercanía y altitud relativa de un PAM sobre indicadores de salud en niños menores de 5 años**

VARIABLES	Hemoglobina	Anemia	HA (OMS)	WA (OMS)	WH (OMS)	Desnutrido	Bajo peso
<b>Cercanía: 5 kilómetros</b>							
Cerca x Pam_alto	-0.0914 (0.0826)	1.879 (3.351)	<b>-0.121*</b> (0.0633)	<b>-0.110*</b> (0.0648)	-0.0406 (0.0659)	3.531 (2.574)	0.0121 (1.393)
Constante	10.13*** (0.0763)	72.82*** (3.147)	-1.093*** (0.0614)	-0.188*** (0.0629)	0.538*** (0.0636)	17.72*** (2.416)	3.165** (1.274)
R2	0.226	0.168	0.236	0.207	0.097	0.113	0.039
<b>Cercanía: 10 kilómetros</b>							
Cerca x Pam_alto	-0.00620 (0.0667)	1.477 (2.736)	<b>-0.136***</b> (0.0525)	<b>-0.0981*</b> (0.0535)	-0.00338 (0.0540)	<b>5.170**</b> (2.152)	1.010 (1.156)
Constante	10.15*** (0.0776)	73.58*** (3.221)	-1.126*** (0.0629)	-0.211*** (0.0646)	0.539*** (0.0649)	19.12*** (2.502)	3.803*** (1.370)
R2	0.226	0.168	0.237	0.207	0.097	0.114	0.039
<b>Cercanía: 15 kilómetros</b>							
Cerca x Pam_alto	0.0465 (0.0681)	-0.811 (2.782)	-0.0839 (0.0536)	-0.0656 (0.0533)	-0.00206 (0.0533)	3.534 (2.264)	0.537 (1.189)
Constante	10.17*** (0.0815)	72.23*** (3.330)	-1.105*** (0.0655)	-0.202*** (0.0667)	0.535*** (0.0666)	18.81*** (2.618)	3.492** (1.446)
R2	0.226	0.168	0.236	0.207	0.097	0.114	0.039
Observaciones	12,116	12,116	12,955	12,955	12,955	12,955	12,955
Controles Niño-Madre	SI	SI	SI	SI	SI	SI	SI
Controles pendientes	SI	SI	SI	SI	SI	SI	SI
Ef. Fijo Provincia	SI	SI	SI	SI	SI	SI	SI

**Nota:** Errores estándares robustos entre paréntesis. \*\*\* 99% sig \*\* 95% sig \* 90 % sig. Los controles niño-madre incluyen edad y sexo del niño, y edad, educación, condición laboral, condición conyugal, condición de jefe de hogar de la madre del niño. Los controles pendientes incluyen la pendiente del centro poblado y del PAM cercano.

Como se mencionó, la validez de los resultados descansa en que estos sean robustos a especificaciones que consideren una posible correlación espacial. La Tabla 4 muestra la especificación que corrige la correlación espacial siguiendo a Conley (1999)<sup>4</sup>. Los resultados dejan de ser significativos. Por otro lado, la Tabla 5 muestra los resultados considerando la segunda forma de corregir la correlación espacial: errores agrupados. Bajo este último se mantienen los resultados encontrados en la especificación principal, aunque con menor significancia.

Tabla 4.

Indicadores de Salud. Errores estándar corregidos por correlación espacial - Conley (1999)

VARIABLES	Hemoglobina	Anemia	HA (OMS)	WA (OMS)	WH (OMS)	Desnutrido	Bajo Peso
<b>Cercanía: 5 kilómetros</b>							
Cerca x Pam_alto	-0.073	2.734	-0.156	-0.088	0.021	2.105	0.344
Err. Est	0.139	4.820	0.184	0.145	0.104	3.467	1.668
<b>Cercanía: 10 kilómetros</b>							
Cerca x Pam_alto	-0.084	3.727	-0.104	0.004	0.113	3.487	0.148
Err. Est	0.122	4.431	0.137	0.112	0.084	2.724	1.175
<b>Cercanía: 15 kilómetros</b>							
Cerca x Pam_alto	0.033	-1.156	-0.064	0.036	0.126	2.904	-0.112
Err. Est	0.096	3.406	0.118	0.091	0.070	2.973	1.407
Observaciones	12,116	12,116	12,955	12,955	12,955	12,955	12,955

**Nota:** \*\*\* 99% sig \*\* 95% sig \* 90% sig. Los controles niño-madre incluyen edad y sexo del niño, y edad, educación, condición laboral, condición conyugal, condición de jefe de hogar de la madre del niño. Los controles pendientes incluyen la pendiente del centro poblado y del PAM cercano.

<sup>4</sup> El método de Conley (1999) requiere gran intensidad computacional, por lo que la especificación propuesta se diferencia de la tabla principal al no incluir efectos fijos. Como el propósito es probar la influencia de la correlación espacial, no creemos que se generen complicaciones.



Tabla 5  
Indicadores de salud. Errores estándar agrupados a nivel distrito

VARIABLES	Hemoglobina	Anemia	HA (OMS)	WA (OMS)	WH (OMS)	Desnutrido	Bajo Peso
<b>Cercanía: 5 kilómetros</b>							
Cerca x Pam_alto	-0.0914 (0.105)	1.879 (3.649)	-0.121 (0.0793)	<b>-0.110*</b> (0.0656)	-0.0406 (0.0594)	3.531 (2.941)	0.0121 (1.620)
Constante	10.13*** (0.0814)	72.82*** (3.208)	-1.093*** (0.0734)	-0.188*** (0.0688)	0.538*** (0.0668)	17.72*** (2.713)	3.165** (1.356)
R2	0.226	0.168	0.236	0.207	0.097	0.113	0.039
<b>Cercanía: 10 kilómetros</b>							
Cerca x Pam_alto	-0.00620 (0.0810)	1.477 (3.107)	<b>-0.136**</b> (0.0623)	<b>-0.0981*</b> (0.0555)	-0.00338 (0.0533)	<b>5.170**</b> (2.377)	1.010 (1.216)
Constante	10.15*** (0.0855)	73.58*** (3.357)	-1.126*** (0.0765)	-0.211*** (0.0700)	0.539*** (0.0673)	19.12*** (2.832)	3.803*** (1.441)
R2	0.226	0.168	0.237	0.207	0.097	0.114	0.039
<b>Cercanía: 15 kilómetros</b>							
Cerca x Pam_alto	0.0465 (0.0816)	-0.811 (3.295)	-0.0839 (0.0637)	-0.0656 (0.0567)	-0.00206 (0.0585)	3.534 (2.523)	0.537 (1.279)
Constante	10.17*** (0.0890)	72.23*** (3.537)	-1.105*** (0.0799)	-0.202*** (0.0736)	0.535*** (0.0727)	18.81*** (3.076)	3.492** (1.584)
R2	0.226	0.168	0.236	0.207	0.097	0.114	0.039
Observaciones	12,116	12,116	12,955	12,955	12,955	12,955	12,955
Controles niño-madre	SI	SI	SI	SI	SI	SI	SI
Controles pendientes	SI	SI	SI	SI	SI	SI	SI
Ef. Fijo departamento	SI	SI	SI	SI	SI	SI	SI
Errores agrup. distrito	SI	SI	SI	SI	SI	SI	SI

**Nota:** Errores estándar robustos en paréntesis. \*\*\* 99% sig \*\* 95% sig \* 90% sig. Los controles niño-madre incluyen edad y sexo del niño, y edad, educación, condición laboral, condición conyugal, condición de jefe de hogar de la madre del niño. Los controles pendientes incluyen la pendiente del centro poblado y del PAM cercano.

Tener resultados significativos que agrupan errores, pero no considerar la corrección de Conley (1999), sugiere tomar los resultados con cuidado. El estudio más cercano con el cual discutir resultados es el de Goltz y Barnwal (2013), quienes encuentran que la cercanía de las unidades mineras aumenta la anemia y la desnutrición; sin embargo, no se evalúa si sus resultados son robustos a correlación espacial siguiendo a Conley (1999), pero sí cuando agrupan errores. Dadas las tablas mostradas, solo podremos decir que se tiene evidencia parcial del efecto de la cercanía de un PAM en indicadores de salud en niños menores de 5 años.

## **Indicadores de producción agropecuaria**

La segunda manera de aproximarnos a los efectos de la contaminación de un PAM sobre el bienestar es analizando la producción agrícola y pecuaria de hogares productores agropecuarios cercanos a estas fuentes de contaminación. Particularmente, nos concentramos en cuatro indicadores bajo un enfoque de función de producción siguiendo a Aragón y Rud (2015), a fin de estimar el impacto en la productividad. Los indicadores son el valor de producción agrícola, el valor de producción pecuario, el número de ganado y la producción agropecuaria. Seguimos las definiciones sobre valor de producción de Escobal y Armas (2015), quienes corrigen valores atípicos en los precios de cada cultivo.

El modelo principal incluye los factores de producción y variables de control utilizadas en Aragón y Rud (2015). Particularmente, se consideran los factores de producción tierra (superficie agropecuaria), trabajo (familiar y asalariado) e insumos intermedios (gasto). Se consideran como variables de control características del jefe de hogar-productor: sexo, edad, edad al cuadrado, estado civil, nivel educativo y si es oriundo del distrito donde vive; características de la unidad agropecuaria: porcentaje de superficie con título de propiedad y superficie bajo riego; y, el grado de pendiente del centro poblado y del PAM cercano.

En la Tabla 6 mostramos los resultados para las definiciones de 'cercanía' de 5, 10 y 15 kilómetros de distancia entre centro poblado y PAM. Al igual que el caso de salud, el coeficiente de interés es aquel sobre la interacción entre la definición de cercanía y la condición de que el PAM se encuentre a mayor altura que el centro poblado. Restringimos la muestra a aquellos centros poblados ubicados a menos de 100 kilómetros de un PAM y no consideramos Lima Metropolitana, por ser un caso particular.

Tabla 6.  
**Resultados principales de cercanía y altitud relativa de un PAM  
sobre indicadores de producción agropecuaria**

VARIABLES	Producción agrícola	Producción pecuaria	Número de ganado	Producción agropecuaria
<b>Cercanía: 5 kilómetros</b>				
Cerca x Pam_alto	-0.131 (0.0975)	-0.104 (0.207)	-0.196 (0.133)	-0.0400 (0.0748)
Constante	3.439*** (0.184)	1.897*** (0.347)	-0.440** (0.217)	3.467*** (0.161)
R2	0.570	0.226	0.265	0.622
<b>Cercanía: 10 kilómetros</b>				
Cerca x Pam_alto	0.103 (0.0762)	-0.140 (0.151)	<b>-0.232**</b> (0.0964)	0.00940 (0.0531)
Constante	3.505*** (0.185)	1.868*** (0.349)	-0.467** (0.217)	3.481*** (0.162)
R2	0.571	0.226	0.266	0.622
<b>Cercanía: 15 kilómetros</b>				
Cerca x Pam_alto	0.0512 (0.0651)	-0.0612 (0.139)	<b>-0.159*</b> (0.0882)	0.0343 (0.0483)
Constante	3.491*** (0.186)	1.897*** (0.352)	-0.457** (0.219)	3.493*** (0.162)
R2	0.570	0.226	0.266	0.622
Observaciones	9,532	8,328	8,328	9,671
Controles JH-UA	SI	SI	SI	SI
Controles pendientes	SI	SI	SI	SI
Ef. Fijo provincia	SI	SI	SI	SI

**Nota:** Errores estándar robustos en paréntesis. \*\*\* 99% sig \*\* 95% sig \* 90% sig. Los controles JH-UA incluyen sexo, edad, estado civil, nivel educativo y distrito de nacimiento del jefe de hogar productor, y porcentaje de tierra con título y porcentaje con riego de la unidad agropecuaria. Los controles pendientes incluyen la pendiente del centro poblado y del PAM cercano.

No encontramos efectos significativos en la productividad agrícola, pecuaria y agropecuaria, pero sí era esperable un valor negativo. Por otro lado, encontramos que el número de ganado es menor en unidades agropecuarias que se hallan relativamente cercanas considerando las distancias de 10 y 15 kilómetros. Las tablas 7 y 8 muestran los resultados al controlar una posible correlación espacial siguiendo a Conley (1999) y agrupando errores, respectivamente. Al agrupar errores el resultado para el número de ganado se mantiene a la distancia de 10 kilómetros; mientras que de la manera alternativa es significativa la menor cantidad de ganado en unidades agropecuarias cercanas a un PAM, en las tres distancias señaladas.

Tabla 7.

**Indicadores de producción agropecuaria. Errores estándar corregidos por correlación espacial - Conley (1999)**

VARIABLES	Producción agrícola	Producción pecuaria	Número de ganado	Producción agropecuaria
<b>Cercanía: 5 kilómetros</b>				
Cerca x Pam_alto	-0.115	-0.106	<b>-0.397**</b>	-0.060
Err. Est	0.115	0.233	0.171	0.079
<b>Cercanía: 10 kilómetros</b>				
Cerca x Pam_alto	-0.049	-0.190	<b>-0.464***</b>	-0.070
Err. Est	0.118	0.170	0.152	0.069
<b>Cercanía: 15 kilómetros</b>				
Cerca x Pam_alto	-0.143	0.007	<b>-0.280*</b>	-0.025
Err. Est	0.101	0.211	0.162	0.064
Observaciones	9,532	8,328	8,328	9,671

**Nota:** \*\*\* 99% sig \*\* 95% sig \* 90 % sig. Los controles JH-UA incluyen sexo, edad, estado civil, nivel educativo y distrito de nacimiento del jefe de hogar productor, y porcentaje de tierra con título y porcentaje con riego de la unidad agropecuaria. Los controles pendientes incluyen la pendiente del centro poblado y del PAM cercano.

Tabla 8  
Indicadores de producción agropecuaria. Errores estándar agrupados

VARIABLES	Producción agrícola	Producción pecuaria	Número de ganado	Producción agropecuaria
<b>Cercanía: 5 kilómetros</b>				
Cerca x Pam_alto	-0.131 (0.114)	-0.104 (0.237)	-0.196 (0.164)	-0.0400 (0.0871)
Constante	3.439*** (0.208)	1.897*** (0.352)	-0.440* (0.226)	3.467*** (0.184)
R2	0.570	0.226	0.265	0.622
<b>Cercanía: 10 kilómetros</b>				
Cerca x Pam_alto	0.103 (0.107)	-0.140 (0.172)	<b>-0.232*</b> (0.122)	0.00940 (0.0656)
Constante	3.505*** (0.199)	1.868*** (0.353)	-0.467** (0.227)	3.481*** (0.183)
R2	0.571	0.226	0.266	0.622
<b>Cercanía: 15 kilómetros</b>				
Cerca x Pam_alto	0.0512 (0.0924)	-0.0612 (0.160)	-0.159 (0.112)	0.0343 (0.0568)
Constante	3.491*** (0.198)	1.897*** (0.362)	-0.457** (0.228)	3.493*** (0.183)
R2	0.570	0.226	0.266	0.622
Observaciones	9,532	8,328	8,328	9,671
Controles JH-UA	SI	SI	SI	SI
Controles pendientes	SI	SI	SI	SI
Ef. Fijo Departamento	SI	SI	SI	SI
Errores Agrup. Distrito	SI	SI	SI	SI

**Nota:** Errores estándar robustos en paréntesis. \*\*\* 99% sig \*\* 95% sig \* 90 % sig. Los controles JH-UA incluyen sexo, edad, estado civil, nivel educativo y distrito de nacimiento del jefe de hogar productor, y porcentaje de tierra con título y porcentaje con riego de la unidad agropecuaria. Los controles pendientes incluyen la pendiente del centro poblado y del PAM cercano.

A diferencia de Aragón y Rud (2015), no se determina que el encontrarse cerca de fuentes contaminantes tenga un efecto significativo en la productividad agrícola. Hay coincidencia con Del Pozo y Paucarmayta (2015) respecto al menor número de ganado. Ambos trabajos agrupan errores, mas no presentan los errores estándar corregidos siguiendo a Conley (1999). Los resultados mostrados permiten afirmar que si bien no se tiene evidencia sobre el efecto en productividad, si se tiene en menor número de ganado.

## Conclusiones

Este estudio es una aproximación exploratoria sobre los efectos de la contaminación minera en el bienestar de poblaciones locales cercanas. Desde las ciencias sociales, la investigación sobre las actividades extractivas suele girar en torno al concepto de la maldición de los recursos naturales, a fin de verificar si esta teoría se cumple en contextos determinados y para variables de empleo e ingresos. Presentamos un concepto más amplio sobre la maldición de los recursos naturales y sostenemos que en un mismo espacio geográfico nacional pueden existir bendiciones y maldiciones, de manera no determinística.

En particular, nos concentramos en los potenciales efectos que los PAM pueden tener en la salud y la producción agropecuaria de las localidades cercanas a estos. Mediante la combinación de elementos de análisis espacial (utilizados en geografía) y análisis econométrico, se obtiene algunos resultados para discutir sobre la cercanía a pasivos ambientales mineros: a) los niños menores de 5 años tienen menor altura y peso según edad, así como mayor probabilidad de encontrarse desnutridos, aunque los resultados son parcialmente robustos; y b) los hogares productores agropecuarios tienen menor número de ganado, resultado robusto a correlación espacial.

Los estudios previos sobre el tema son pocos y no suelen considerar que, dado el carácter de los fenómenos estudiados, es necesario contrastar los resultados con métodos o especificaciones que consideren correlación espacial. Las limitaciones de datos no permiten un análisis más profundo como el de considerar la antigüedad de los pasivos y su nivel de contaminación. En conjunto, debido al carácter exploratorio del estudio y a las limitaciones de la información disponible, interpretamos los resultados como correlaciones más no en términos de causalidad. Instamos a la audiencia especializada a continuar produciendo investigación sobre los efectos de la contaminación minera.

## Referencias bibliográficas

- Alexeev, M. y Conrad, R. (2009). *The elusive curse of oil*. *The Review of Economics and Statistics*, 91 (3), 586-598.
- Aragón, F., Chuhan-Pole, P. y Land, B. (2014). *The local economic impacts of resource abundance*. Manuscrito no publicado.
- Aragón, F. y Rud, J. (2013). *Natural resources and local communities: evidence from a peruvian gold mine*. *American Economic Journal: Economic Policy*, 5 (2), 1-25.
- Aragón, F. y Rud, J. (2015) *Polluting industries and agricultural productivity: evidence from mining in Ghana*. *The Economic Journal*, 126, 1980-2011.
- Auty, R. (1993) *Sustaining Development in Mineral Economies: The Resource Curse Thesis*. London: Routledge.

- Bebbington, A. (Ed.) (2012). *Social conflict, economic development and extractive industry: evidence from South America*. London: Routledge.
- Brollo, F., Nannicini, T., Perotti, R. y Tabellini, G. (2013). The political resource curse. *The American Economic Review*, 103 (5), 1759-1796
- Caselli, F. y Michaels, G. (2013). Do oil windfalls improve living standards? Evidence from Brazil. *American Economic Journal: Applied Economics*, 5 (1), 208-238.
- Collier, D., Hoeffler, A. (2005) Resource Rents, Governance and Conflict. *The Journal of Conflict Resolution*. 49 (4), pp. 625-633.
- Corden, W. y Neary, J. (1982). Booming sector and de-industrialisation in a small open economy. *The Economic Journal*, 92, 825–848.
- Conley, T. (1999). GMM estimation with cross sectional dependence. *Journal of Econometrics*. 92 (1), pp. 1-45.
- Cust, J. y Poelhekke, S. (2015). The local impacts of natural resource extraction. *Annual Review of Resource Economics*, 7, 251-268.
- Deaton, A. y Cartwright, N. (2016). Understanding and misunderstanding randomized controlled trials. National Bureau of Economic Research, working Paper no. 22595.
- Del Pozo, C., y Paucarmayta, V. (2014). *Impacto de la minería en la producción agropecuaria en el Perú, impactos heterogéneos y determinación de los canales de transmisión*. Lima: CIES - Concurso Anual de Investigación.
- Dell, M. (2010) The persistent effects of Peru's Mining Mita. *Econometrica*, 78 (6), pp. 1863-1903.
- Dell, M., Jones, B. y Olken, B. (2014) What do we learn from weather? The New Climate-Economy Literature. *Journal of Economic Literature*. Vol 52 (3) 740-798.
- Dufflo, E. y Pande, R. (2005) Dams. MIT Working Paper.
- Easterly, W. y Rose L., (2003) Tropics, Germs, and Crops: The Role of Endowments in Economic Development. *Journal of Monetary Economics*, 50(1), January 2003.
- Emel, J., Plisinski, J. y Rogan. J. (2014) Monitoring geomorphic and hydrologic change at mine sites using satellite imagery: The Geita Gold Mine in Tanzania. *Applied Geography* 54, pp. 243-249
- Ghose, M. (2010). Impact of mining on the female community: A perspective of female miners in the Indian context. *Minerals and Energy - Raw Materials Report* 19: 16-24.

- Glave, M. y Kuramoto, J. (2002) *Minería, minerales y desarrollo sustentable en el Perú*. Lima: GRADE.
- Graff-Zivin, J. y Neidell, M. Environment, Health, and Human Capital. *Journal of Economic Literature* 51, 689–730.
- Greenstone, M. Hornbeck, R. y Moretti, E. (2010) Identifying Agglomeration Spillovers: Evidence from Winners and Losers of Large Plant Openings. *Journal of Political Economy*, 118 (3), 536–598.
- Hanna, R. y Oliva, P. (2011). The Effect of Pollution on Labor Supply: Evidence from a Natural Experiment in Mexico City. Technical Report, National Bureau of Economic Research.
- Hendrychova, M. y Kabrna, M. (2016) An analysis of 200-year-long changes in landscape affected by large-scale surface coal mining: History, present and future. *Applied Geography*, 74, pp. 151-159.
- Herrera, P. y Millones, O. (2012) Aproximando el costo de la contaminación minera sobre los recursos hídricos: metodologías paramétricas y no paramétricas. *Economía* 35 (70) pp. 9-59.
- Hsiang, S. (2016) *Climate Econometrics*. NBER Working Paper.
- Kotsadam, A. y Tolonen, A. (2014) African Mining, Gender, and Local Employment. *World Development* 83, 325-339.
- Leite, C., y Weidman, J. “Does Mother Nature Corrupt? Natural Resources, Corruption, and Economic Growth,” IMF Working Paper, 1999, (99/85).
- León, J., Benavides, M. y otros (2016) *Los efectos de la violencia doméstica sobre la salud infantil de los niños y niñas menores de cinco años en el Perú*. Lima: GRADE: Documento de Trabajo 82.
- Loayza, N. y Rigolini, J. (2016) *The Local Impact of Mining on Poverty and Inequality: Evidence from the Commodity boom in Peru*. *World Development*. Elsevier, vol. 84 (C), 219-234.
- Marshall, F., Ashmore, M. y Hinchcliffe, F. (1997) A Hidden Threat to Food Production: Air Pollution and Agriculture in the Developing World. International Institute for Environment and Development London.
- Mcculloch, J. (2003). Women Mining Asbestos in South Africa, 1893-1980. *Journal of Southern African Studies* 29(2): 413-432.
- Menz, F. y Seip, H. (2004) Acid Rain in Europe and the United States: an Update. *Environmental Science & Policy* 7 (4), 253–265.



- Michaels, G. (2011) The Long Term Consequences of Resource-Based Specialisation. *The Economic Journal*, 121 (551), 31–57.
- Miller, J., (1988) Effects on Photosynthesis, Carbon Allocation, and Plant Growth Associated with Air Pollutant Stress,. En “Assessment of Crop Loss from Air Pollutants,” Springer, 287–314.
- Moretti, E. (2011) Local Labor Markets. *Handbook of Labor Economics*, 4, 1237-1313.
- Orihuela, J. (2017) Institutions and Place: Bringing Context Back into the Study of Resource Curse. *Journal of Institutional Economics*, 1-24, doi: 10.1017/S1744137417000236.
- Phelps, N., Atienza, M. y Arias, M. (2015). Encore the enclave: the changing nature of the industry enclave with illustrations from the mining industry. *Economic Geography*, 91 (2), pp 119-146.
- Phillips, J. (2016) Climate Change and Surface Mining: A review of environment-human interaction and their spatial dynamics. *Applied Geography*, 74, pp. 95-108.
- Rau, T., Reyes, L. y Urzua, S. (2015) Early Exposure to Hazardous Waste and Academic Achievement: Evidence from a Case of Environmental Negligence. *Journal of Association of Environmental and Resource Economists* 2 (4) pp. 527-563.
- Ross, M. (2013) *The Oil Curse: How Petroleum Wealth Shapes the Development of Nations*. Princeton University Press.
- Sachs, J. y Warner, A. (1995). Natural resource abundance and economic growth. National Bureau of Economic Research, Reporte técnico.
- Sachs, J. y Warner, A. (2001). The curse of natural resources. *European Economic Review*, 45 (4), 827-838.
- Ticci, E. y Escobal, J. (2015) Extractive industries and local development in the Peruvian Highlands. En *Environment and Development Economics*, 18 (6),1-26.
- Thorp, R., Battistelli, S., Guichaoua, Y., Orihuela, J. y Paredes, M. (2012). *The developmental challenges of mining and oil: lessons from Africa and Latin America*. Basingstoke: Palgrave McMillan
- Van Hoecke, E. (2006). The invisible work of women in the small mines of Bolivia. En *Women Miners in Developing Countries: Pit Women and Others*. Burlington, VT: Ashgate.
- Van der Ploeg, F. (2011). Natural resources: curse or blessing?. *Journal of Economic Literature*, 49 (2), 366-420

Valencia Areli (2016) Human Rights Trade-off in Times of Economic Growth. The Long-Term Capability Impacts of Extractive.Led Development. Palgrave.

WHO - World Health Organization- (2017) *Inheriting a sustainable world? Atlas on children's health and the environment*. Suiza: WHO.



