



# MANUAL METODOLÓGICO

## VIEME Encuesta de Microemprendimiento 2019



---

**DEPARTAMENTO DE ESTADÍSTICAS DEL TRABAJO**

**SUBDEPARTAMENTO DE ESTADÍSTICAS ESTRUCTURALES DEL TRABAJO**

**DEPARTAMENTO DE METODOLOGÍA E INNOVACIÓN ESTADÍSTICA**

---

**Manual Metodológico VI Encuesta de Microemprendimiento**

**Marzo 2020**

**Instituto Nacional de Estadísticas**

El uso de un lenguaje que no discrimine ni marque diferencias entre hombres y mujeres ha sido una preocupación en la elaboración de este documento. Sin embargo, y con el fin de evitar la sobrecarga gráfica que supondría utilizar en castellano “o/a” para marcar la existencia de ambos sexos, se ha optado por utilizar –en la mayor parte de los casos– el masculino genérico, en el entendido de que todas las menciones en tal género representan siempre a hombres y mujeres, abarcando claramente ambos sexos.

Dirección Nacional: Morandé N°801, piso 22, Santiago, Chile.

Fono: 232461010

Fax: 562 267 12 169

Código Postal: 8340148

Sitio web: [www.ine.cl](http://www.ine.cl)

Correo electrónico: [ine@ine.cl](mailto:ine@ine.cl)

Santiago de Chile

## Índice

1. INTRODUCCIÓN.....	8
2. ANTECEDENTES GENERALES DE LA ENCUESTA.....	10
2.1. Antecedentes del producto estadístico .....	11
2.2. Objetivos del producto estadístico .....	12
2.2.1. Objetivo general .....	12
2.2.2. Objetivos específicos de la encuesta.....	12
3. MARCO DE REFERENCIA .....	14
3.1. Marco conceptual .....	14
3.1.1. Marco conceptual emprendedores.....	14
3.1.2. Economía informal.....	15
3.2. Referencias nacionales e internacionales .....	20
3.2.1. La contribución de la Economía informal al PIB: El caso peruano .....	22
4. DISEÑO ESTADÍSTICO.....	26
4.1. Características generales del diseño muestral.....	26
4.1.1. Objetivo del diseño muestral.....	26
4.1.2. Población objetivo .....	26
4.1.3. Unidades estadísticas.....	27
4.1.4. Método de muestreo.....	27
4.2. Características del Marco Muestral .....	27
4.2.1. Cobertura del marco muestral .....	28
4.2.2. Estratificación del marco muestral .....	28
4.2.3. Depuración del listado de trabajadores independientes.....	29
4.3. Cálculo y distribución de la muestra.....	30
4.3.1. Cálculo del tamaño muestral .....	30
4.3.2. Distribución de la muestra a nivel regional.....	31
4.3.3. Distribución de la muestra por submuestras .....	34
4.4. Selección de unidades muestrales .....	35
4.4.1. Selección de viviendas .....	35
4.4.2. Selección de informantes .....	36
5. METODOLOGÍA DE RECOLECCIÓN DE DATOS.....	37
5.1. Técnica y Métodos de recolección de datos.....	37
5.1.1. Principales técnicas de recolección de datos.....	37
5.1.2. Medio de recolección de información utilizado.....	37
5.1.3. Período de referencia .....	37
5.1.4. Fechas de recolección de datos y procesos complementarios .....	38

5.1.5.	Identificar el tipo de entrevista y la confirmación del equipo de trabajo (encuestadores, supervisores y proporción por región).....	39
5.2.	Diseño del instrumento de captura de información.....	40
5.2.1.	Actualización cuestionario Piloto.....	41
5.2.2.	Diseño prueba piloto.....	43
5.2.3.	Levantamiento y análisis prueba piloto .....	46
5.2.4.	Definición cuestionario oficial.....	46
5.3.	Prueba y evaluación del instrumento .....	48
5.3.1.	Entrevistas cognitivas .....	48
5.3.2.	Observaciones Técnicas .....	50
5.3.3.	Mesas de conversación .....	51
5.3.4.	Análisis Base de datos .....	53
5.4.	Método de capacitación para la recolección de datos oficial .....	54
5.4.1.	Capacitaciones encuestadores y supervisores .....	54
5.4.2.	Capacitación analistas.....	55
5.5.	Supervisión.....	55
5.5.1.	Supervisión directa .....	55
5.5.2.	Supervisión indirecta .....	55
6.	METODOLOGÍA DE PROCESAMIENTO DE DATOS.....	57
6.1.	Codificación y clasificación .....	57
6.1.1.	Codificación CAENES y CIUO.....	57
6.1.2.	Codificación y reclasificación especifique .....	58
6.2.	Análisis de consistencia y edición.....	59
6.2.1.	Diseño e implementación de validaciones .....	59
6.2.2.	Revisión valores atípicos .....	60
6.3.	Imputación y creación de variables .....	64
6.3.1.	Depuración y estimación de gastos, ingresos y ganancias .....	64
6.3.2.	Construcción de variable ganancia final.....	69
6.3.3.	Visualización base de datos .....	70
6.3.4.	Generación variables trabajo no remunerado .....	70
6.4.	Cálculo y ajuste de ponderadores .....	72
6.4.1.	Cálculo ponderador base.....	72
6.4.2.	Ajuste por no respuesta (y suavizamiento) .....	74
6.4.3.	Calibración .....	75
7.	METODOLOGÍA DE ANÁLISIS .....	77
7.1.	Análisis descriptivo.....	77

7.2.	Comparación con fuentes externas.....	77
8.	ALCANCE DE LA PRESENTACIÓN DE RESULTADOS.....	79
8.1.	Productos estadísticos para publicar .....	79
8.1.1.	Publicaciones y anuarios.....	79
8.1.2.	Metodología.....	79
8.1.3.	Manuales.....	79
8.1.4.	Formularios.....	79
8.1.5.	Bases de datos .....	79
8.1.6.	Metadatos .....	80
8.2.	Difusión de los resultados .....	80
8.3.	Accesibilidad a los datos y marco legal .....	80
8.3.1.	Marco legal .....	80
8.3.2.	Proceso Anonimización Base de datos.....	81
9.	GLOSARIO.....	86
10.	BIBLIOGRAFÍA.....	97
11.	ANEXOS .....	100
11.1.	Anexo 1: análisis del número óptimo de donantes .....	100
11.2.	Anexo 2: perfiles de donantes .....	103

## INDICE DE TABLAS

Tabla 1: Sectores institucionales del SCN 2008.....	17
Tabla 2: Unidades y actividades de producción .....	20
Tabla 3: Experiencias Internacionales en la aplicación de Encuestas a Micronegocios ...	21
Tabla 4: Cuenta de producción del Sector Informal y sector institucional de los hogares Perú, (Miles de Millones de nuevos soles corrientes.....	24
Tabla 5: Criterios utilizados en el cálculo del tamaño muestral .....	30
<i>Tabla 6: Errores esperados asociados al parámetro de interés según tamaño muestral objetivo .....</i>	<i>33</i>
Tabla 7: Total de viviendas seleccionadas según región y mes de levantamiento .....	34
Tabla 8: Errores observados asociados al parámetro de interés según número de viviendas que responden .....	35
Tabla 9: Encuestadores y supervisores por región .....	40
Tabla 10: Cambios cuestionario Prueba Piloto VI EME.....	42
Tabla 11: Distribución de muestra agrícola por región y trimestre de referencia .....	45
Tabla 12: Prueba Piloto. Distribución de viviendas muestra final según región y trimestre de referencia.....	45
<i>Tabla 13: Distribución de muestra según región y sector de actividad .....</i>	<i>45</i>
Tabla 14: Duración de entrevistas según Formato.....	47
Tabla 15: cambios cuestionario Prueba Piloto VI EME .....	47
Tabla 16: Preguntas y Conceptos medidos en Entrevistas Cognitivas.....	49
Tabla 17: número de hallazgos por cada tipo .....	51
Tabla 18: Comparación de codificación ENE- EME .....	58
Tabla 19: casos totales y casos outliers.....	62
Tabla 20: Medias para casos en rango y casos outliers.....	63
Tabla 21: Ausencia de respuesta en preguntas de gastos, ingresos y ganancias.....	65
Tabla 22: Listado de Actividades de Trabajo No Remunerado consultados VI EME .....	70
Tabla 23: Medidas de distribución horas semanales destinadas al trabajo no remunerado versión original.....	71
Tabla 24: Medidas de distribución horas semanales destinadas al trabajo no remunerado versión truncada .....	72
Tabla 25: Total de unidades elegibles, que responde y tasa de respuesta .....	75
Tabla 26. Total de microemprendedores estimados a partir de la ENE – período MAM 2019 .....	76
Tabla 27: Variables clave y número de categorías.....	84
Tabla 28: Cumplimiento del K-anonimato (nivel individual) . <b>¡Error! Marcador no definido.</b>	
Tabla 29: Cumplimiento del K-anonimato después de depuración (nivel individual) . <b>¡Error! Marcador no definido.</b>	

## **INDICE DE FIGURAS**

Figura 1: Relación entre Economía no observada y el Sector Informal .....	19
Figura 2: Las fuentes de información de la cuenta satélite.....	23
Figura 3: Pauta de evaluación indirecta. ....	56
Figura 4: Proceso de búsqueda de donantes.....	68

## **INDICE DE DIAGRAMAS**

Diagrama 1: Línea de tiempo de trabajo de recolección de datos.....	38
Diagrama 2: Esquema de participación de los informantes en la ENE y EME. ....	39
Diagrama 3: Etapa de diseño de materiales .....	40
Diagrama 4: Proceso construcción cuestionario final.....	46
Diagrama 5: Proceso pregunta-respuesta.....	48
Diagrama 6: Proceso de codificación .....	57
Diagrama 7: Implementación métodos SDC y evaluación del riesgo .....	85

## 1. INTRODUCCIÓN

El Instituto Nacional de Estadísticas (INE) en el marco del convenio suscrito con el Ministerio de Economía, fomento y turismo, tiene a cargo desde el año 2013, la producción de la Encuesta de Microemprendimiento (EME), la que, en conjunto con otras fuentes de información, generan estadísticas e indicadores relevantes sobre las unidades económicas de menor tamaño.

En este sentido, el objetivo principal de la encuesta es lograr una caracterización completa de los microemprendimientos que se desarrollan a nivel nacional y regional, permitiendo identificar las formas de funcionamiento de estas unidades, sus niveles de ingresos, el empleo generado y la situación de formalidad, entre otros datos.

Esta encuesta es una herramienta de enorme valor estadístico para el país, puesto que es el único estudio de este tipo que se realiza a lo largo de todo Chile, abarcando unidades económicas pequeñas, ya sea formales o informales, pertenecientes a todos los sectores económicos.

A diferencia de otras encuestas a empresas o negocios, la EME es una encuesta a hogares de segunda fase<sup>1</sup>, cuya población objetivo son los trabajadores por cuenta propia o empleadores dueños de una microempresa con hasta 10 trabajadores incluyendo a su dueño. Esta encuesta se denomina como “Encuesta mixta” ya que permite obtener información de dos unidades de análisis el Microempresario y su negocio.

El proceso estadístico de la última versión de la encuesta se expone al público a través del presente este manual, para quienes buscan conocer y comprender su funcionamiento. El documento se estructura en base a los siguientes contenidos:

- El capítulo I y II, introducen y exponen los antecedentes de la encuesta y los objetivos que tiene este producto estadístico, enfocándose principalmente en la caracterización que permite realizar de las unidades económicas informales.
- Luego, en el capítulo III se detalla el marco teórico desde donde se obtienen los criterios que dan sustento al estudio. Además, se incluye experiencia de otros países en esta medición.
- El capítulo IV presenta el diseño muestral de la encuesta, el cual expone las características generales del diseño muestral, el marco muestral y la cobertura de este, el cálculo de la muestra y su distribución, para finalmente detallar la metodología de selección de unidades: viviendas e informantes.
- En su capítulo V se desarrolla la metodología de recolección de datos en levantamiento oficial. Describe la prueba piloto realizada para diseñar el cuestionario del levantamiento oficial y las técnicas utilizadas para evaluar los resultados de esta etapa. Además, se detalla el método de la capacitación y de la supervisión del trabajo en terreno.
- El capítulo VI de metodología de procesamiento de datos, se ahonda en la

---

<sup>1</sup> La EME es aplicada a las personas ocupadas clasificadas como trabajadores por cuenta propia o empleadores dueños de unidades económicas con hasta 10 trabajadores a través de la Encuesta Nacional de Empleo durante el trimestre marzo -abril -mayo



identificación y corrección de inconsistencias, la clasificación y codificación de variables y la generación de variables adicionales. Junto con, exponer la metodología de edición e imputación de las preguntas de gastos, ingresos y ganancias; y el cálculo final de los ponderadores.

- En el siguiente capítulo (VII), metodologías de análisis, se describen los métodos estadísticos utilizados para analizar e interpretar los datos.
- Finalmente, en el último capítulo (VIII), se exponen los productos estadísticos producidos y cómo acceder a ellos, junto con de indicar las limitaciones de la información entregada y la confidencialidad de los datos.

## **2. ANTECEDENTES GENERALES DE LA ENCUESTA.**

El Ministerio de Economía, Fomento y Turismo, en su tarea por diseñar e implementar políticas públicas que promuevan la competitividad y productividad de la economía, realiza un constante esfuerzo por generar información con estándares de calidad sobre el emprendimiento en Chile, sus determinantes y las características de estas fuentes productivas en el país. Para ello, desde el 2013 le ha encomendado al Instituto Nacional de Estadísticas, el diseño, recolección, procesamiento y análisis de la Encuesta de Microemprendimiento, uno de sus principales instrumentos de aproximación a la realidad de los microemprendedores en Chile.

El fomento productivo, la integración económica y la modernización de la estructura productiva del país requieren contar con información actualizada y completa. Para esto, existen distintas fuentes de información, tales como registros administrativos y encuestas a establecimientos, las cuales permiten caracterizar a una parte importante de las unidades económicas del país. No obstante, estas fuentes no logran capturar la información completa de las microempresas, pues se basan en listados de empresas formales que excluyen a la parte de las unidades económicas que pertenecen al sector informal.

En ese sentido, la Encuesta de Microemprendimiento (EME) es un instrumento de enorme valor pues permite llegar a todos los microemprendimientos ya sean formales o informales. En particular, resulta relevante la medición que realiza la encuesta respecto al sector de la economía con características de informalidad, puesto que es actualmente el único estudio de este tipo que se realiza a lo largo de todo Chile y a través de todos los sectores económicos.

En cuanto a la caracterización de la población ocupada, los trabajadores independientes representan un grupo de interés relevante dentro de la estructura ocupacional del país. Según datos de la Encuesta Nacional de Empleo el porcentaje de trabajadores independientes sobre el total de ocupados fue de 26,4% en el trimestre marzo-abril-mayo, además, la mayor parte de ellos (97,6%), corresponde a Microemprendedores. La EME permite realizar una caracterización profunda de este tipo de trabajadores entregando información relevante sobre la motivación por la que iniciaron su emprendimiento, la fuente de financiamiento inicial, el período de funcionamiento, entre otras variables que no son posibles de obtener por medio de otra fuente de información.

## 2.1. Antecedentes del producto estadístico

La Encuesta de Microemprendimiento (EME) tiene sus inicios en el año 2006 a partir de una consultoría sobre microempresas informales que permitió realizar una primera caracterización de negocios no registrados ante las autoridades tributarias, los cuales no son visibilizadas por las encuestas tradicionales a empresas.

En el año 2009 se llevó a cabo la primera versión de la Encuesta de Microemprendimiento, que entregó una caracterización de las microempresas a lo largo del país. La segunda versión fue ejecutada por el Centro de Microdatos en el año 2011.

Debido a la distancia entre el período de aplicación de la primera y segunda fase de las encuestas, las dos primeras versiones presentaron altas tasas de no respuesta y elevados cambios en la situación ocupacional entre los informantes. Para poder solucionar este problema, a partir del 2013, el INE asume el proceso estadístico de la encuesta, desarrollando desde el diseño del cuestionario hasta la publicación de los datos.

Entre los diversos cambios implementados por el INE el año 2013, se redujeron considerablemente los contenidos del cuestionario, pasando de un total de 200 preguntas y una hora de duración de la entrevista, a 68 preguntas con tiempo de entrevista de aproximadamente 22 minutos. Esto permitió reducir el rechazo y desgaste de los informantes, mejorando las tasas de respuesta y logro de la encuesta.

Entre las modificaciones al cuestionario, desataca la creación del módulo E “Situación de Formalidad del negocio”, el que permite identificar diferentes grados de informalidad de las unidades económicas y aportar así con insumos para la generación de políticas públicas que incentiven la incorporación al sector formal. Para ello, se incluyeron 11 preguntas, que permitieron realizar un primer análisis de informalidad laboral en base a los lineamientos establecidos por la Organización Internacional del Trabajo (OIT).

Con el fin de enfocar el análisis de la encuesta sólo a los microemprendimientos, en la V EME (2017) se modificó la población objetivo del estudio de trabajadores independientes a microemprendedores, es decir, trabajadores independientes dueños de una unidad económica con hasta 10 empleados o trabajadores (incluyendo al propio dueño), lo que contribuyó con mejorar la precisión de las estimaciones<sup>2</sup> y, al mismo tiempo, vincular el nombre del estudio más estrechamente con la medición que efectivamente se realiza.

El cambio de la población objetivo facultó la incorporación de una grilla de trabajadores, que da la posibilidad de aproximarse a las principales características del personal empleado por la persona microemprendedora. La estructura de la información recopilada se repite

---

<sup>2</sup> Producto de que la EME es una encuesta mixta que se ancla a una encuesta de hogares como es la Encuesta Nacional de Empleo (ENE), no es posible estimar de manera precisa el número de unidades económicas de mayor tamaño, las que típicamente son obtenidas a través de encuestas a establecimientos cuya muestra es extraída del directorio de empresas. Esto ocurre debido a que las viviendas donde residen los dueños de estas unidades económicas de unidades de mayor tamaño (mediana o grandes) tienen menor probabilidad de ser seleccionadas en la EME, ya que el marco de muestreo proveniente de la ENE tiene limitantes de acceso al estrato socioeconómico alto donde comúnmente residen estos empresarios.

para cada uno de los empleos generados por el informante, por ello desde la V versión, se decidió publicar la información de estas preguntas en una base de datos independiente, en la que cada fila representa la información de un empleado y no de un microemprendedor.

En la versión actual de la EME, se incorporaron preguntas para capturar con mayor detalle la información de las ganancias de las unidades económicas, el valor agregado de éstas y los salarios de los trabajadores. Además, se incluyó un módulo de trabajo no remunerado que permite conocer el tiempo destinado a las actividades de trabajo de cuidado y trabajo doméstico no remunerado.

## **2.2. Objetivos del producto estadístico**

### **2.2.1. Objetivo general**

Caracterizar los microemprendimientos que se desarrollan a nivel nacional, permitiendo conocer las limitantes y los elementos facilitadores que tienen las unidades económicas de menor tamaño, para llevar a cabo sus actividades dentro del mercado laboral, a partir de una muestra representativa de viviendas particulares que contienen al menos a una persona microemprendedora, a nivel nacional y regional.

### **2.2.2. Objetivos específicos de la encuesta**

La VI Encuesta de Microemprendimiento permite caracterizar distintas condiciones de los Microemprendedores y sus negocios las temáticas abordadas a través del cuestionario se resumen en las siguientes categorías:

- 1) Estudiar el grado de formalidad de los micronegocios:** la principal virtud de la encuesta es que permite caracterizar a los microemprendimientos pertenecientes al sector informal de la economía. Para ello, el módulo E contiene una serie de preguntas enfocadas a detallar la situación de formalidad de los negocios y sus dueños.
- 2) Estudiar la productividad y los ingresos de las microempresas:** parte importante de las brechas de los salarios o ganancias obtenidas por los trabajadores o dueños se explican por las brechas en productividad de los micronegocios, en este sentido la encuesta considera una serie de preguntas que permiten calcular el valor agregado de los micronegocios y con ello aproximarse a la estimación de algunas medidas de productividad.
- 3) Estimar la ocupación generada por las microempresas y las características de esta:** algunos de los aspectos que permite estudiar la EME es la existencia de ocupación ocasional o las relaciones entre los empleados y los dueños, cómo se distribuye la carga laboral y si los trabajadores tienen un contrato con garantías legales.
- 4) Motivación emprendedora:** identificar el principal motivo porque las personas deciden emprender si es por necesidad, oportunidad o bien una tradición familiar.
- 5) Observar las formas en las que operan y funcionan las unidades económicas:** identificar la escala de funcionamiento de los negocios y estudiar

la organización y planificación del trabajo de los microemprendedores, como también los factores utilizados en el proceso productivo y la capacidad de poder realizar una separación de ellos.

- 6) **Medir y caracterizar el acceso al sistema financiero:** identificar las fuentes de financiamiento de los microemprendedores y aproximarse al nivel de deuda de los micronegocios.
- 7) **Calcular la magnitud del trabajo no remunerado:** calcular la cantidad de tiempo destinado al trabajo no remunerado por sexo e identificar la relación de este con otras variables que permitan explicar de mejor manera las dinámicas laborales de los microemprendedores.

### 3. MARCO DE REFERENCIA

El siguiente apartado contiene el marco conceptual que sustenta la encuesta de Microemprendimiento, el cual se basa principalmente en la medición de la economía informal.

#### 3.1. Marco conceptual

El marco conceptual en el cual se circunscribe la EME tiene dos objetos de caracterización principal, la de los microemprendedores y sus negocios y, en particular, el segmento que pertenece al sector informal. De esta manera en este subapartado se detalla, en primer lugar, los conceptos de microemprendedores y su lugar en la fuerza de trabajo en el caso de Chile, y en segundo lugar, se detalla los conceptos de vinculados con la economía informal.

##### 3.1.1. Marco conceptual emprendedores

Según lo establecido en la resolución de la vigésima Conferencia Internacional de Estadísticas del trabajo organizada por la OIT “*Los emprendedores son trabajadores que crean trabajo para sí mismos y potencialmente para otros por medio del establecimiento y operación de una empresa*” (OIT, 2020). Para poder identificar a este grupo, se establece a la categoría ocupacional de los trabajadores independientes, como la categoría más adecuada para identificar a los emprendedores y recopilar estadísticas al respecto.

Esta misma resolución señala que, es necesaria información adicional del contexto nacional para proveer estadísticas completas sobre el emprendimiento. En el caso chileno, las personas ocupadas pueden ser clasificadas en dos grandes categorías ocupacionales: los trabajadores dependientes y los trabajadores independientes. Los primeros, son trabajadores que tienen dependencia y subordinación a una unidad económica (o empleador) y que entregan su fuerza de trabajo a cambio de un pago en dinero o especies. En contraste, los trabajadores independientes, son trabajadores sin dependencia ni subordinación que ejercen su profesión, oficio o negocio.

A su vez, los trabajadores independientes se dividen en dos grupos, los empleadores y los trabajadores por cuenta propia. Los empleadores, son aquellos trabajadores independientes que emplean al menos un trabajador de forma permanente. A su vez, los trabajadores por cuenta propia son trabajadores independientes que no generan empleo de forma permanente, pero pueden trabajar con socios o emplear a trabajadores de forma esporádica.

En el caso del mercado laboral chileno y en particular para el estudio de esta encuesta, se decidió restringir el análisis a un subconjunto de los emprendedores o trabajadores independientes, que son los que trabajan en una unidad económica de hasta 10 trabajadores denominada como Microempresa. Debido a posibles problemas en la captura de emprendedores de alto impacto y que estos son estudiados principalmente a partir de una encuesta a establecimientos.

Según lo establecido en la resolución de la vigésima Conferencia Internacional de Estadísticas del trabajo organizada por la OIT “*los emprendedores son trabajadores que crean trabajo para sí mismos y potencialmente para otros por medio del establecimiento y operación de una empresa*” (OIT,2002). Para poder identificar a este grupo se establece a la categoría ocupacional de los trabajadores independientes como la categoría más adecuada para identificar a los emprendedores y recopilar estadísticas al respecto. Esta misma resolución señala que es necesaria información adicional del contexto nacional para proveer estadísticas completas sobre el emprendimiento.

Los trabajadores independientes se dividen en dos grupos. Los empleadores son trabajadores independientes que emplean al menos un trabajador de forma permanente; y los trabajadores por cuenta propia son trabajadores independientes que no generan empleo de forma permanente, sin embargo, pueden trabajar con socios o emplear a trabajadores de forma esporádica.

En el caso del mercado laboral chileno, y en particular para el estudio de esta encuesta se decidió restringir el análisis a un subconjunto de los emprendedores o trabajadores independientes, que son los que trabajan en una unidad económica de hasta 10 trabajadores denominada como Microempresa. El motivo de esto es la baja captura de emprendedores de alto impacto y la mayor heterogeneidad de este grupo que en muchos ámbitos no son comunes a los empleadores de menor tamaño y trabajadores por cuenta propia.

### 3.1.2. Economía informal

Según la OIT (2002) la economía informal corresponde al conjunto de “todas las actividades económicas de los trabajadores y las unidades económicas que están- en la legislación o en la práctica- no cubiertas o insuficientemente contempladas por los sistemas formales”<sup>3</sup>.

La relevancia de la economía informal no se concentra solamente en la cantidad de ocupación asociada a ella (ocupación informal y ocupación en el sector informal) sino también en el aporte del sector informal a la economía. Esta área de la economía no observada es un lugar importante de producción de bienes y servicios, y representa, especialmente para los países en desarrollo, una significativa parte de sus transacciones económicas.

En el sector informal se llevan a cabo actividades económicas que van desde la venta callejera hasta las prestaciones de servicios profesionales de la más variada índole, sin la declaración ante las autoridades tributarias nacionales. Estas se desarrollan en viviendas con o sin una instalación especial, en la vía pública o en establecimientos. Y generalmente, son llevadas a cabo por personas que trabajan por cuenta propia o empleadores que contratan un número escaso de asalariados, o sólo cuentan con trabajadores familiares que les prestan apoyo sin percibir una remuneración. Este amplio abanico de características conjuga al sector informal y constituyen las razones por las cuales muchas veces, las

---

<sup>3</sup> 90ª Reunión de la Conferencia Internacional del Trabajo (2002).

actividades desarrolladas dentro de sus límites son subvaloradas e ignoradas por los métodos estadísticos tradicionales.

Una medición exhaustiva de la economía informal les permite a los países cuantificar la participación de la ocupación informal (en el sector formal y en el sector informal) dentro del número total de ocupados y, además, estimar la contribución del sector informal a sus economías. Accediendo, así, a la posibilidad de aclarar las zonas oscuras que escapan de las mediciones estadísticas tradicionales y con ello, a una visión integrada de las actividades económicas desarrolladas en el país. De este modo, se puede contar con insumos más sólidos para entender la dinámica del mercado laboral, enfocar las políticas públicas y mejorar las condiciones laborales de las personas con ocupaciones informales.

Para las cuentas nacionales, “la población total ocupada (sea formal o no) interviene en la creación de valor medido por el PIB, que es la suma del valor agregado por todas las unidades de producción, de mercado o no, que participan en la economía nacional” (OIT, 2013, p.277), por lo tanto, en aquellas economías donde el aporte del sector informal al PIB es de gran importancia, resulta necesario establecer mediciones detalladas de las dimensiones y cualidades de dicho sector. De modo que, para el fenómeno de la informalidad, la medición de su aporte al PIB y al mercado en general, mediante parámetros coherentes con los del Sistema de Cuentas Nacionales (SCN), constituye el camino hacia la visibilización de su importancia dentro de las economías nacionales.

En la actualidad, a pesar de que la producción estadística de algunos países brinda datos importantes para la caracterización de la informalidad, estos carecen de los lineamientos metodológicos necesarios para evidenciar su participación en la económica nacional (Séruzier, 2010). A su vez, el SCN no cuenta con las herramientas para acceder a una medición directa de la informalidad, por lo que no tiene luces de las características individuales de las unidades económicas que la componen, ni del impacto de las medidas nacionales para mejorar las temáticas relacionadas con el fenómeno. En este escenario, la vinculación de estas dos áreas beneficia el cumplimiento de sus objetivos.

#### I. El Sistema de Cuentas Nacionales (SCN 2008).

El Sistema de Cuentas Nacionales (SCN) es un conjunto de recomendaciones aceptadas internacionalmente, para la elaboración de “mediciones de la actividad económica de acuerdo con convenciones contables estrictas, basadas en principios económicos” (SCN, 2008, p. 1). Este marco estadístico, brinda cuentas macroeconómicas completas, comparables, coherentes, integradas y flexibles “para ser utilizadas en la elaboración de políticas, el análisis y la investigación” (OIT, 2013, p.277).

El SCN proporciona una fotografía de la economía en un momento dado, pero al ser elaboradas de forma sucesiva, concede flujos de información básicos para el análisis de la situación económica y la elaboración de medidas para el desarrollo. En él, se configuran ciertas magnitudes e indicadores como el Producto Interno Bruto (PIB) y se pueden observar las interacciones y el comportamiento de los sectores económicos que lo componen.



El SCN tiene dos tipos de categorías de Unidades Institucionales o Sujetos de las Transacciones, fundamentales para la medición del consumo, la producción y la acumulación: los hogares y las entidades jurídicas. Su principal característica es que se trata de unidades económicas con la capacidad de tener activos e incurrir en pasivos a su nombre, pudiendo, además, llevar a cabo transacciones con otras unidades.

De forma particular el SCN, agrupa a las Unidades Institucionales en cuatro sectores mutuamente excluyentes, definidos debido a “sus funciones principales, comportamiento y objetivos” (OIT, 2012, p. 278). Dichas unidades son presentadas en el cuadro 1:

**Tabla 1: Sectores institucionales del SCN 2008**

SECTORES INSTITUCIONALES DEL SCN						
Gobierno	Sociedades (Financieras y no financieras)	Hogares				Instituciones sin fines de lucro al servicio de los hogares
		Hogares con empresas no constituidas en sociedad que se encuentran registradas y/o tiene más de un número dado de asalariados (pertenecientes al sector informal)	Hogares con empresas no constituidas en sociedad, que no están registradas y/o tienen menos de un número dado de asalariados ( <b>Empresas del sector informal</b> ).		Hogares con empresas no constituidas en sociedad que producen exclusivamente para uso final propio.	Hogares institucionales como prisiones, órdenes religiosas y casa de retiro para personas mayores, etc.
			Empresas informales por cuenta propia (sin asalariados permanentes).	Empresas de empleadores informales		

Fuente: Elaboración propia a partir del SCN, 2008

En este punto, resulta necesario precisar que al utilizar que el concepto de “sector informal”, no es el mismo que el concepto de sector utilizado por el SCN, el cual divide los “Sectores” según la naturaleza de ellos. Como se evidencia en la tabla 1 el sector informal corresponde a un subconjunto del sector institucional de los hogares, haciendo exclusiva alusión a los establecimientos productivos pertenecientes a los hogares no conformados en sociedad, excluyendo a las cuasi sociedades<sup>4</sup>. Por lo tanto, para una eventual recopilación de las cuentas del sector hogares, sería necesario acceder a la información de todos sus subconjuntos.

En el SCN, el sector informal es entendido como todas aquellas unidades de producción de bienes y servicios, pertenecientes al sector institucional hogares, que existen

<sup>4</sup>Según el sistema de cuentas nacionales, “algunas empresas no constituidas en sociedad funcionan en todo –o casi en todo- como si fueran sociedades. Estas empresas se denominan cuasi-sociedades en el SCN y se incluyen con las sociedades en los “sectores de las sociedades no financieras y de las sociedades financieras” (ONU 2009, inciso 4.42).

primordialmente para la generación de ocupación e ingresos y no se encuentran constituidas en sociedad, excluidas las cuasi-sociedades.

El Sistema de Cuentas Nacionales (2008), aborda a través de uno de sus capítulos, aspectos relacionados con la economía informal. En este, se pone de manifiesto el desafío que existe en todas las economías: lograr desarrollar una captura integra de las actividades económicas que en ellas se desarrollan.

## II. Sector informal y economía no observada

La Economía no Observada (ENO), corresponde a aquella sección de difícil acceso, principalmente, por el carácter ilegal de los bienes y actividades en ella desarrolladas. En la ENO existen dos tipos de producciones: la producción ilegal y la producción oculta o subterránea.

La producción ilegal, se desarrolla dentro de los límites del SCN y comprende a la producción de bienes o servicios cuya comercialización, tenencia y distribución es penada por la ley. Por ejemplo: la venta o producción de drogas.

Por otro lado, la producción oculta o subterránea, abarca aquellas actividades dedicadas a la producción de bienes o prestación de servicios legales, que se desarrollan fuera del conocimiento de las autoridades pertinentes, principalmente, para evadir el pago de los impuestos asociados, el pago de prestaciones sociales (propias o a trabajadores) o el cumplimiento de ciertas normativas legales, procedimientos administrativos y registros fiscales o estadísticos. En resumen, es aquella economía que escapa de todo mecanismo de visibilización normativa.

Existen coincidencias entre los criterios de la ENO y del sector informal, no obstante, como podemos observar en la figura 1, estas se superponen solo en un punto: aquellas actividades no observadas que se llevan a cabo de manera informal.

**Figura 1: Relación entre economía no observada y el sector informal**



Fuente: adaptación figura 25.1 SCN, 2008.

Los distintos grupos ejemplificados en la figura 1 son posible de encontrar en muchos países, donde gran parte de las empresas del sector informal pueden tener algún tipo de registro o pagar impuestos o permisos, no cumplan todas las exigencias administrativas y legales, es decir, que pertenezcan al sector observado e informal (círculo discontinuo). Por

otra parte, algunos segmentos sustanciales de la producción subterránea provienen de empresas pertenecientes al sector formal (círculo continuo). Algunos ejemplos incluyen la producción de bienes y servicios ‘no declarada’ o ‘al negro’, transacciones financieras o ingresos de la propiedad no declarados, sobrestimación de los gastos deducibles de los impuestos, la ocupación de trabajadores y salarios no declarados y horas extraordinarias no visibilizadas de los asalariados declarados.

**Tabla 2: Unidades y actividades de producción**

Unidades de producción	Actividades de producción		
	Legal		Ilegal
	No subterránea	Subterránea	
Empresas del sector formal			
Empresas del sector informal (a)			
Hogares (b)			

*Nota: (a) Según la definición de la resolución de la 15 CIET (con exclusión de los hogares que emplean trabajadores domésticos remunerados). (b) Hogares que producen bienes exclusivamente para uso final propio y hogares que emplean trabajadores domésticos asalariados.*  
Fuente: OIT 2013.

Según se indica en la tabla 2 “cualquier tipo de unidad de producción (empresas del sector formal, empresas del sector informal, hogares) puede llevar a cabo cualquier tipo de actividad legal y no subterránea, legal y subterránea e ilegal” (OIT, 2013, p. 54). De forma particular, gran parte de las actividades y servicios desarrollados en el marco del sector informal, son perfectamente legales y su origen no está necesariamente impulsado por la intención de evadir los pagos o legislación correspondiente. Algunas de estas unidades económicas, ostentan un nivel de producción y capital tan bajo, que dar cumplimiento a cualquier normativa o pago de impuestos, resulta inviable.

### 3.2. Referencias nacionales e internacionales

En la región existen algunas experiencias que tienen como objetivo analizar y caracterizar a las microempresas pertenecientes al sector institucional de los hogares y/ o al sector informal. Para ello, lo más recomendado es realizar una encuesta mixta, las cuales son denominadas de esta manera porque permiten obtener información del hogar o entrevistado y de su empresa, teniendo una doble unidad de análisis.

A su vez las encuestas mixtas se pueden dividir en dos tipos: encuestas mixtas modulares y encuestas mixtas independientes. Las encuestas mixtas modulares incorporan un módulo específico que permite tener información de las unidades económicas o empresas, el cual es aplicado en el mismo levantamiento de la encuesta a hogares o de fuerza de trabajo. En cambio, la encuesta mixta independiente implica el levantamiento de una encuesta adicional sobre las características de la empresa o negocio, a partir de una subpoblación

obtenida desde la aplicación de una encuesta en primera fase, el cual se transforma en el marco muestral de la encuesta en segunda fase.

Además, la experiencia regional sobre la aplicación de este tipo de encuestas es diferente en cada país. De hecho, existen distintas definiciones de la población objetivo y en consecuencia del diseño y del marco muestral. Algunos de estos países distinguen la población objetivo a través de un número de asalariados por micronegocio, en otros casos, es a partir del sector al cual pertenece la unidad económica (si es un hogar o bien si la unidad económica de la cual es dueño el entrevistado pertenece al Sector Informal). A continuación, se muestra en la tabla 3 el resumen de las experiencias presentes en el levantamiento de estas encuestas en la región:

**Tabla 3: Experiencias internacionales en la aplicación de encuestas a micronegocios**

N°	País	Nombre de Encuesta	Tipo de Encuesta	Población objetivo: Trabajadores independientes dueños de una...	Fecha 1° Publicación	Fecha última publicación
1	Brasil	Encuesta Economía Informal Urbana	Independiente	Unidad Económica no agrícolas con menos de 5 asalariados	1998	2003
2	Costa Rica	Encuesta Nacional de Microempresas de Hogares	Independiente	Unidad Económica perteneciente al Sector de los hogares	2012	2020
3	Colombia	Gran Encuesta Integrada de Hogares	Modular	Unidad Económica de hasta 9 asalariados	2000	2018
4	México	Encuesta Nacional de Micronegocios	Independiente	Unidad Económica no agrícolas de hasta 10 o 15 asalariados	1993	2012
5	Perú- INEI	Encuesta Nacional de Hogares	Modular	Unidad Económica del Sector Informal	2012	2017

Fuente: INE, 2019.

Una de las experiencias más longevas corresponde a la Encuesta de Micronegocios de México (ENAMIN) realizada desde 1992 con una periodicidad bianual. La ENAMIN es una encuesta en segunda fase aplicada a una subpoblación de la Encuesta Nacional de Ocupación y Empleo (ENOE), la cual abarca todas las actividades no agrícolas, formales o informales. Esta encuesta, realiza un límite por número de trabajadores según la actividad económica para las ramas de comercio, transporte, construcción y servicios, incluye unidades con seis o menos trabajadores (incluidos el propietario o director); para la manufactura, incluye las unidades con 16 o menos trabajadores (incluido el jefe).

La segunda experiencia más antigua sobre el levantamiento de este tipo de encuestas en la región es la Encuesta de Economía Informal Urbana (ECINF) realizada por el Instituto Brasileño de Geografía y Estadísticas (IBGE de Brasil). Si bien esta encuesta fue pensada para levantarse cada cinco años, sólo fue aplicada en dos oportunidades los años 1997 y 2003. La ECINF fue una encuesta mixta independiente, sin embargo, la particularidad de

ella es que la primera fase no corresponde a una Encuesta propiamente tal, sino a un proceso de enumeración de manzanas, donde dentro de las manzanas seleccionadas se encuentran las viviendas que contienen trabajadores independientes dueños de una unidad no agrícola con menos de 5 asalariados, los cuales pasan a formar parte del marco muestral de la encuesta.

La Encuesta de Microestablecimiento aplicada desde el 2001 por el DANE (Colombia) ha sufrido una serie de ajustes metodológicos para mejorar la comparabilidad y confiabilidad de sus estimaciones. Por ello, el año 2012 se optó por realizar un rediseño de la investigación a través de la implementación de dos instrumentos: un módulo de Micronegocios en la Gran Encuesta Integrada de Hogares y la encuesta a Microestablecimientos tipo panel. La población objetivo de esta investigación corresponde a los trabajadores independientes dueños de una unidad económica con hasta 9 asalariados, sin importar el sector al que pertenece este negocio.

Otra de las encuestas destacadas en la región corresponde a la Encuesta Nacional de Microempresas de Hogares (ENAMEH) de Costa Rica. La ENAMEH está dirigida a pequeños negocios pertenecientes a los hogares. Se trata de una encuesta independiente, donde la primera fase corresponde a la Encuesta Nacional de Hogares (ENAHO), desde la que se extrae el marco muestral de la ENAMEH. Una de sus principales utilidades, es que entrega estadísticas básicas requeridas para la estimación de la cuenta del Sector institucional de los hogares que realiza el Banco Central de Costa Rica (BCCR).

Finalmente, una de las experiencias más recientes y completa en la aplicación de este tipo de encuestas es la de Instituto Nacional de Estadística e Informática (INEI) de Perú, quienes, a partir de un módulo integrado en la Encuesta Nacional de Hogares aplicado solo a trabajadores independientes del sector informal, logran estimar la contribución de este sector al PIB. A continuación, se abordará en detalles esta experiencia:

### 3.2.1. La contribución de la Economía informal al PIB: El caso peruano

Actualmente Perú es de las pocas naciones a nivel latinoamericano que miden la contribución de la producción informal al PIB. A través de un módulo aplicado en la Encuesta Nacional de Hogares capturar el valor agregado de la producción de las unidades económicas pertenecientes al sector informal y la ocupación informal de la economía. A partir de estos resultados, el Instituto Nacional de Estadísticas e Informática (INEI) realiza una estimación que permite mantener la sincronía entre las cuentas de producción, de ingresos y de gastos.

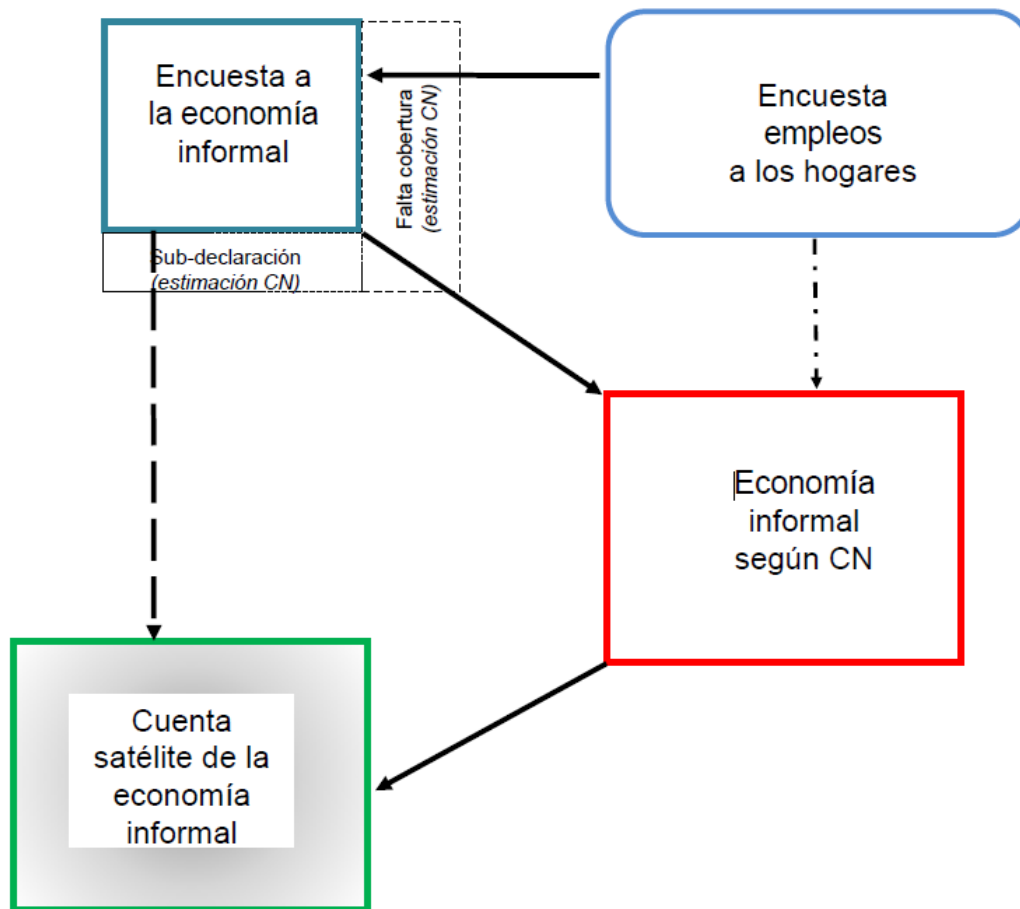
Para cuantificar la contribución del sector informal al PIB el INEI cuenta con distintas fuentes que posibilitan tener un panorama completo del impacto de la economía informal en la Contabilidad Nacional (CN). De esta manera, el módulo de sector informal anexo a la Encuesta Nacional de Hogares (ENAHO) tiene como finalidad convertirse en un insumo para el cálculo de la contribución del sector informal al PIB. Esta información luego es ajustada a partir de cuentas nacionales, las cuales pueden tener diferencias por problemas de cobertura y/o subdeclaración de la encuesta (limitaciones propias de toda encuesta a

hogares) y por otros errores no muestrales, producto (por ejemplo) de que el informante idóneo puede no conocer las respuestas a ciertas preguntas necesarias para ser clasificado en el sector informal.

Para cuantificar el aporte final de la economía informal al PIB es necesario comparar y buscar la convergencia en 3 campos:

- i. Cuentas de producción por actividad, diferenciadas según segmentos de unidades productivas (valor agregado, consumo intermedio, etc.) y sectores institucionales,
- ii. una matriz de ocupación por categorías utilizando los mismos criterios (actividad, segmentos de producción y sectores), y
- iii. una ubicación aproximada de los establecimientos informales presentes en estos segmentos de producción, a partir del seguimiento de la ocupación.

**Figura 2: Las fuentes de información de la cuenta satélite**



Fuente: Sérurier (2011b, p.74).

El primer paso para cuantificar el aporte del sector informal al PIB es identificar las unidades económicas del sector informal en la economía, el cual se sitúa en el sector institucional de los hogares y está compuesto por una parte de las unidades de producción de este sector. Estas unidades son identificadas por medio de la segunda fase aplicada en la ENAHO. A

partir de los datos de esta medición y la cuadratura con las otras fuentes de información, el sector informal en la Economía el año 2007 se encuentra detallado en el siguiente cuadro.

**Tabla 4: Cuenta de producción del sector informal y sector institucional de los hogares Perú, (miles de millones de nuevos soles corrientes)**

Tipo de Unidad Económica	Valor Bruto Producción	Consumo Intermedio	Valor Agregado Bruto	Remuneraciones	Otros Impuestos	Exced. De Explot/Ingreso Mixto
<b>Sector institucional de los hogares</b>	<b>156,5</b>	<b>61,7</b>	<b>94,8</b>	<b>18,4</b>	<b>0,0</b>	<b>76,4</b>
Sector "formal" de los hogares <sup>a</sup>	41,2	20,2	21,1	5,7	0,0	15,3
Sector informal	100,5	40,7	59,8	9,7	0,0	50,1
<i>Agropecuario</i>	26,5	7,4	19,0	3,4	0,0	15,6
<i>No agropecuario</i>	74,0	33,3	40,7	6,2	0,0	34,5
Hogares con producción para uso final propio	14,8	0,9	14,0	3,0	0,0	10,9
<b>Otros sectores institucionales</b>	<b>407,6</b>	<b>209,2</b>	<b>198,4</b>	<b>79,8</b>	<b>1,5</b>	<b>117,2</b>
<b>Total Economía</b>	<b>564,1</b>	<b>270,9</b>	<b>293,2</b>	<b>98,1</b>	<b>1,5</b>	<b>193,5</b>

Nota: a/ Según la definición adoptada, todo el sector "formal" del sector institucional de los hogares es no agropecuario.  
Fuente: Instituto Nacional de Estadística e Informática – Cuentas nacionales en INEI Perú (2017).

En la tabla 4 se puede observar la cuenta de producción de la economía a distintos niveles de desagregación. Con el objetivo de poder visibilizar el aporte del sector informal dentro de la producción total, la primera desagregación divide el aporte del valor bruto entre el sector institucional de los hogares con el de los "otros sectores institucionales". Luego, la segunda desagregación, corresponde a la división del sector institucional de los hogares en tres grupos: El sector formal del sector institucional de los hogares, el sector informal<sup>5</sup> y los hogares con producción para uso final propio<sup>6</sup>. Además, se tiene una subdivisión dentro del sector informal que diferencia la producción realizada por la actividad agropecuaria de las demás ramas.

A través de esta medición, para el año 2007, podemos apreciar que el sector informal representa alrededor del 20% del Valor Agregado Bruto total de la economía (59,8 sobre 293,2). La cuenta de producción del sector informal no agropecuario presenta diferencias en cuanto a la estructura del sector informal agropecuario. Por ejemplo, en el consumo intermedio, el cual es de mayor importancia en las actividades no agropecuarias (45%) que

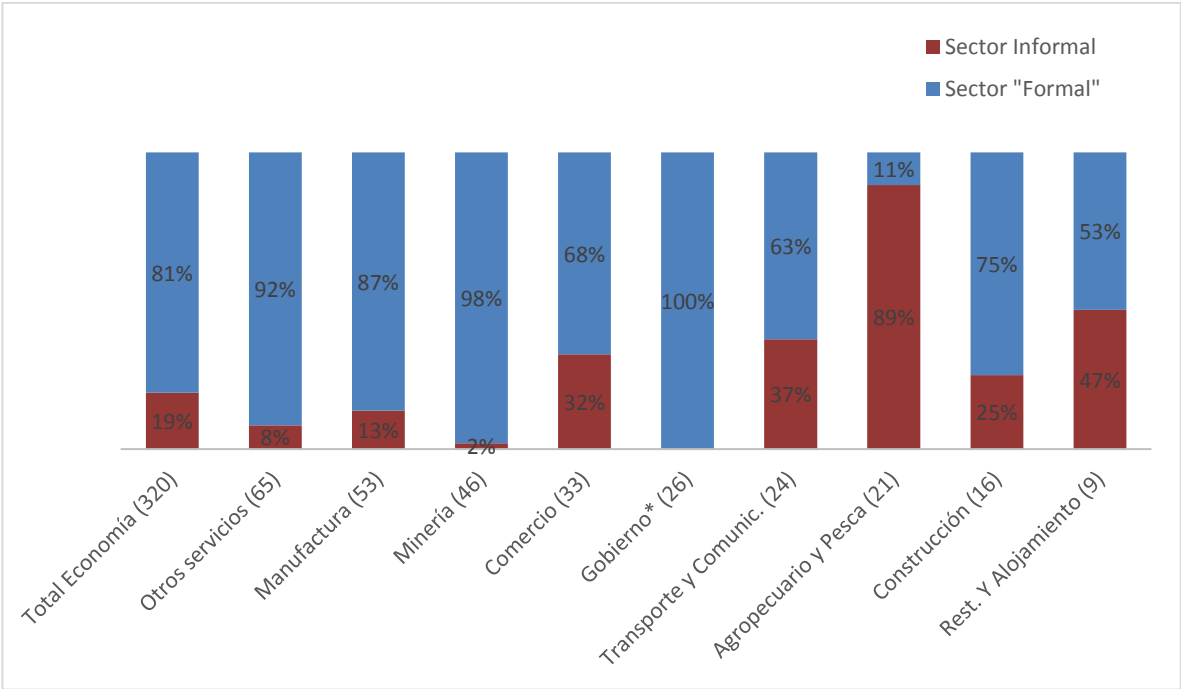
<sup>5</sup> Es importante señalar que por definición el sector "formal" del sector institucional de los hogares no contiene unidades productivas del sector primario, dado que, en la definición adoptada, todas las unidades productivas agropecuarias, mineras y de pesca, del sector de los hogares se han definido como informales.

<sup>6</sup> Los hogares con producción exclusiva para uso final propio no son "formales" ni informales, sino que constituyen un grupo aparte.



en las agropecuarias (28%). Por otra parte, y en línea con el resultado anterior, existe un mayor valor agregado en la producción realizada en la actividad agropecuaria (72%) que en las demás actividades (55%), esto se debe a que los productos de la actividad agropecuaria presentan costos menores, pues, los productos agrícolas al ser bienes primarios no implican una inversión en la transformación de la materia prima o insumos para la elaboración del bien final.

**Gráfico 1: Participación del sector informal en el PBI por actividad económica, Perú 2007 (Miles de millones de nuevos soles corrientes)**



\*Incluye Salud y educación pública  
Fuente: Instituto Nacional de Estadística e Informática – Cuentas nacionales en INEI Perú (2017).

Además, al separar la actividad agropecuaria de las demás ramas de actividad podemos apreciar que el VAB del sector informal no agropecuario representa el 15% del VAB de la economía no agropecuaria, mientras que para la actividad agropecuaria el sector informal representa el 89% del VAB.

## 4. DISEÑO ESTADÍSTICO

El diseño muestral se compone de todos los elementos necesarios para construir una muestra que sea representativa del fenómeno de estudio. Este capítulo explica en detalle estos elementos para la VI Encuesta de Microemprendimiento. En primera instancia se presentan las características del diseño muestral, especificando los objetivos de este, la población objetivo, las unidades estadísticas, y el método de muestreo. Luego se da paso a las características del marco muestral, entre las que se mencionan la estratificación del marco y su cobertura, así como la depuración del listado de individuos. El tercer punto de este capítulo corresponde al cálculo y distribución de la muestra, en el que se expone la forma en que se calculó el tamaño muestral, la distribución regional de la muestra y la composición de las submuestras del estudio. Finalmente, se presenta una descripción de la forma en la que se seleccionaron las unidades muestrales, tanto de las viviendas como de los informantes.

### 4.1. Características generales del diseño muestral

El objetivo principal de la EME es lograr una caracterización profunda de los microemprendimientos que se desarrollan a nivel nacional ya sean formales o informales. Con este fin, se realizó una encuesta a hogares que recogió una muestra representativa de viviendas particulares ocupadas a nivel nacional y regional buscando alcanzar aquellas que contaban con personas con microemprendimientos. La razón por la que se optó por realizar una encuesta en hogares se funda en que éstos permiten observar unidades económicas de menor tamaño, las cuales en muchos casos no están inscritas en el Servicios de Impuestos Internos por lo que pueden quedar fácilmente excluidas de las encuestas tradicionales a establecimientos comerciales.

#### 4.1.1. Objetivo del diseño muestral

El objetivo del diseño muestral fue determinar y distribuir un número de viviendas y microemprendedores que permitan obtener estimaciones del parámetro de interés. En el caso de la VI Encuesta de Microemprendimiento, la variable de interés fue la “Proporción de microemprendedores por cuenta propia sobre el total de microemprendedores” según los niveles de precisión establecidos<sup>7</sup> a nivel nacional y regional.

#### 4.1.2. Población objetivo

La población objetivo está conformada por quienes trabajan por cuenta propia o personas empleadoras dueñas de una microempresa con hasta 10 trabajadores y/o trabajadoras (incluyendo a la o el informante dueño de la unidad económica), denominadas en su conjunto como personas microemprendedoras, que residen en viviendas particulares ocupadas dentro del territorio nacional.

---

<sup>7</sup> Los niveles de precisión son: error absoluto de 1,3% y error relativo de 1,5%, a nivel nacional. Para las regiones, el error absoluto máximo es 8,2% y, el error relativo no supera el 11,0% correspondiente a la Región de Magallanes.

#### 4.1.3. Unidades estadísticas

La unidad de análisis de la VI EME son los microemprendedores, entendidos como trabajadores por cuenta propia y empleadores con hasta 10 trabajadores (incluida la persona dueña de la actividad económica). En tanto la unidad de muestral, es decir la unidad que es objeto del proceso de muestreo, fueron las viviendas en las que habitaban microemprendedores. Esto implicó que el proceso de selección de la muestra identifica viviendas en las que se sabe que reside al menos un microemprendedor, entrevistándose, finalmente, todas las personas susceptibles de ser catalogadas de dicha manera que sea posible encontrar en dicha unidad.

#### 4.1.4. Método de muestreo

El diseño muestral de la VI EME contó de dos fases, siendo la primera el levantamiento de la Encuesta Nacional de Empleo (ENE). Esta encuesta tiene como objetivo caracterizar y cuantificar la población de 15 años y más respecto de su situación en el mercado laboral. El trabajo realizado por la ENE sirvió para identificar a los individuos que compondrán la segunda fase del muestreo de la Encuesta de Microemprendimiento. Esta primera fase del diseño tuvo como población objetivo a toda la población del país residente en viviendas particulares ocupadas y utilizó un muestreo probabilístico, estratificado y bietápico, donde las unidades primarias de muestreo correspondieron a manzanas en el área urbana y secciones en el área RAU<sup>8</sup> - rural; mientras que, las unidades secundarias de muestreo fueron las viviendas particulares ocupadas.<sup>9</sup>

La segunda fase se inicia luego de la realización de trabajo de campo de la ENE. Se analizaron los resultados de esta, clasificándose las viviendas en dos grupos dependiendo de si en éstas habitaba o no, en el período de referencia, al menos un microemprendedor. Las viviendas en las que no se registró la presencia de microemprendedores se descartan, generándose el marco de muestreo con solo aquellas viviendas con presencia de microemprendedores. En particular, la EME ocupó como referencia para generar su marco muestral el levantamiento de marzo, abril y mayo de 2019 de la Encuesta Nacional de Empleo.

## 4.2. **Características del Marco Muestral**

Un marco muestral se define como un conjunto de unidades en que todos sus elementos son inequívocamente identificables, mediante algún procedimiento o listado, a los que es posible asociarles una probabilidad de selección distinta de cero de acuerdo con la estrategia de muestreo. En el caso particular de la VI EME, el marco muestral se encuentra organizado en forma jerárquica, de acuerdo con la división político-administrativa del territorio nacional. Las unidades se ordenan en forma descendente según región, provincia

---

<sup>8</sup> Resto de área urbana. Concepto proveniente del marco muestral de viviendas construido con el Censo 2002.

<sup>9</sup> Cabe señalar que, en el marco del fortalecimiento de la ENE, se definió un nuevo marco y diseño muestral. Para mayor información visitar: <https://www.ine.cl/estadisticas/sociales/mercado-laboral/ocupacion-y-desocupacion>

y comuna y al interior de cada comuna, se conforma la división censal que da origen a las áreas geográficas urbanas y rurales.

#### 4.2.1. Cobertura del marco muestral

La cobertura de un marco muestral corresponde al porcentaje de individuos que contiene el marco sobre el total de individuos de la población que se quiere representar, en este caso todos los trabajadores por cuenta propia y empleadores con hasta 10 trabajadores incluyéndose. Para poder identificar a los elementos de la población en estudio se consideraron individuos tanto de zonas urbanas como rurales, ordenados en dos marcos distintos, de manzanas y de secciones, respectivamente.

La VI EME utilizó un diseño muestral bifásico, por lo tanto, comparte las propiedades de cobertura de dos marcos muestrales, primero el utilizado para la selección de las viviendas de la ENE (correspondiente al período marzo-abril-mayo 2019); y segundo el marco utilizado para la selección de los “microemprendedores”.

El marco muestral del INE, usado como base para la ENE y todas las encuestas de hogares que se levantan en la institución, contempla a la población que reside en viviendas particulares ocupadas, excluyendo a la población que habita en viviendas colectivas, tales como: viviendas con más de 5 hogares, hogares de ancianos, hospitales, cárceles, conventos, entre otros; y a la población que reside en la calle. Sin embargo, este marco sí incluye a los hogares de personas que habitan y trabajan dentro de dichas viviendas colectivas, como porteros, conserjes y otros.

Además, el marco muestral utilizado por la ENE excluye las viviendas ubicadas en las 22 áreas geográficas definidas por el INE como áreas de difícil acceso (ADA) o de alto costo<sup>10</sup>. Por otro lado, para optimizar el trabajo de campo y dadas las características de las unidades muestrales del área urbana (manzanas) se descartaron del marco muestral, previo a la selección de las unidades, las manzanas que contienen 7 o menos viviendas.

#### 4.2.2. Estratificación del marco muestral

La estratificación, previa a la selección de la muestra, corresponde al proceso de agrupar los elementos de la población según su homogeneidad respecto a características determinadas. La agrupación de estos elementos debe cumplir con ciertas características con el fin de mejorar la precisión estadística de las estimaciones: debe contener unidades homogéneas internamente y heterogéneas entre sí, así como debe ser mutuamente excluyentes y colectivamente exhaustivas, es decir, cada elemento en la población debe ser asignado solo a un estrato y ningún elemento de la población puede quedar excluido.

El marco muestral contó con una estratificación geográfica según la división político-administrativa que, en conjunto con la división censal, según Censo 2002, y la rama de actividad prevalente en las unidades primarias de muestreo dio origen a las áreas urbanas y rurales.

---

<sup>10</sup> General Lagos, Colchane, Ollagüe, Juan Fernández, Isla de Pascua, Cochamó, Chaitén, Futaleufú, Hualaihué, Palena, Lago Verde, Guaitecas, O'Higgins, Tortel, Laguna Blanca, Río Verde, San Gregorio, Cabo de Hornos (Ex - Navarino), Antártica, Primavera, Timaukel, Torres del Paine.

La estratificación del marco de la ENE, correspondiente a la primera fase, dio origen a los siguientes estratos:

- **Ciudad o grandes Centros Urbanos (CD):** Conformadas por ciudades o conjuntos de ciudades adyacentes con 40.000 o más habitantes.
- **Resto de Área Urbana (RAU):** Conformadas por conjuntos de Centros Urbanos con menos de 40.000 habitantes.
- **Área Rural (R):** Conformado por el conjunto de entidades clasificadas como rurales de acuerdo con un tamaño poblacional menor a 1.000 habitantes, o entre 1.001 y 2.000 habitantes con predominio de población económicamente activa dedicada a actividades primarias<sup>11</sup>.

En la segunda fase, la cobertura geográfica del marco muestral de la VI EME es similar al de la ENE. Lo anterior implicó que la Encuesta de Microemprendimiento se realiza a lo largo del territorio nacional, estratificándose de acuerdo con las 16 regiones que posee el país.

#### 4.2.3. Depuración del listado de trabajadores independientes

Debido a que, en la primera fase, la ENE fue contestada por un informante idóneo<sup>12</sup> (proxy), es decir, una persona que entrega información respecto a su situación y la de todos los integrantes de su hogar, pudiendo suceder que en un hogar se identifique la presencia de un microempresario, pese a no ser encuestado directamente durante la aplicación de la ENE. Esto lleva a que pueda existir una fuente de error no muestral, denominada error de clasificación que es propio de las encuestas de hogares y ocurre cuando una persona es clasificada como independiente en la ENE pero, que al ser sometido al cuestionario de la EME se determine que no lo es. Esa fue la razón por la cual el marco muestral obtenido a partir de la ENE fue revisado en profundidad para descartar todas las unidades que no cumplen los criterios técnicos para ser clasificadas efectivamente como personas que tienen un microemprendimiento, permitiendo mitigar los problemas posteriores de levantamiento debido a la existencia de casos fuera de muestra (no elegibles). Dentro de los casos que fueron eliminados del listado se encontraban mozos, empaquetadores, personal del servicio doméstico quienes trabajan de forma dependiente para un hogar o establecimiento.

En total, la depuración del marco correspondió a 8,8% de casos descartados por ser potenciales unidades no elegibles<sup>13</sup>, observándose los mayores cambios en la región de Los Lagos (12,2%) y, los menores, en la región de Magallanes (5,6%).

---

<sup>11</sup> Toda aquella actividad relacionada con la extracción de recursos naturales (agricultura, caza, pesca, minería, entre otras).

<sup>12</sup> La ENE define como informante idóneo a toda persona de 15 años o más, miembro del hogar, que reside habitualmente en las viviendas seleccionadas.

<sup>13</sup> En la EME son unidades no elegibles aquellos individuos que en la ENE fueron clasificados como microempresarios, según información proporcionada por informante proxy, sin embargo, al momento de realizar el trabajo de campo se observa que la persona seleccionada no era un microempresario, o también en el caso que el individuo haya cambiado de estado (dejó de ser microempresario).

### 4.3. Cálculo y distribución de la muestra

El tamaño muestral es concebido para obtener estimaciones del parámetro de interés con niveles de precisión establecidos para esta versión, que busca representatividad a nivel nacional y regional.

El primer cálculo del tamaño muestral, en base al cual se programó el trabajo de campo que se desarrolló durante 2019, se realizó con la información de la ENE levantada en el trimestre marzo, abril y mayo de 2018 (MAM 2018).

Se utilizó como parámetro de interés la proporción de microemprendedores por cuenta propia sobre el total de microemprendedores y sus estadísticos asociados.

Se trabajaron escenarios con similares errores de estimación, donde finalmente, se eligió el escenario con un error absoluto de 1,3% y un error relativo de 1,5% a nivel nacional. Con esto, se obtuvo una muestra objetivo de 7.086 viviendas que, considerando la tasa de no respuesta observada en la V EME 2017, dio como resultado un tamaño con sobremuestreo de 8.457 viviendas.

Para obtener los tamaños regionales, se asignó el 70,0% de las viviendas disponibles en el marco de selección, donde reside al menos un microemprendedor, exceptuando las regiones de: Antofagasta, Ñuble y Magallanes, donde se asignó el 85,0% con el objetivo de disminuir el error muestral.

#### 4.3.1. Cálculo del tamaño muestral

En el cálculo del tamaño muestral se usó como parámetro de interés la proporción de microemprendedores por cuenta propia sobre el total de microemprendedores. En la Tabla 5, se presentan definiciones y consideraciones utilizadas en la obtención del tamaño muestral.

**Tabla 5: Criterios utilizados en el cálculo del tamaño muestral**

Parámetro	Descripción
<b>Variable de diseño</b>	$I$ : Variable Bernoulli que considera los siguientes valores $I = \begin{cases} 1; \\ 0; \end{cases}$ Si el microemprendedor es cuenta propia En otro caso
<b>Parámetro asociado</b>	Proporción de microemprendedores por cuenta propia sobre el total de microemprendedores en la región $r$ ( $p_{r0}$ )
<b>Estimador asociado</b>	Estimador de razón: $r = \frac{\text{Estimación del número de microemprendedores por cuenta propia en la región}}{\text{Estimación del número de microemprendedores en la región}}$
<b>Niveles de estimación</b>	Nacional Regional
<b>Errores de muestreo</b>	- Nacional: El error absoluto no debe superar 1,3% y; el error relativo 1,5% - Regional: Los errores absolutos no deben superar 8,2% y; los errores relativos 11,0%

Fuente: Instituto Nacional de Estadísticas, 2019.

#### 4.3.2. Distribución de la muestra a nivel regional

A continuación, se describe en detalle cada una de las etapas realizadas para definir el número de viviendas a encuestar en cada región  $r$ .

Sean:

$r$	: Subíndice que identifica a la región
$p_{r0}$	: Parámetro de interés en la región $r$
$\hat{p}_{r0}$	: Estimación del parámetro de interés en la región $r$
$\bar{m}_{r0}$	: Número promedio de viviendas logradas por unidad primaria de muestreo en la región $r$ en la ENE trimestre MAM 2018
$n_{r0}$	: Número de manzanas o secciones levantadas en trimestre MAM 2018 (ENE) en la región $r$
$n_{r0} \cdot \bar{m}_{r0} = m_{r0}$	: Número de viviendas logradas en trimestre MAM 2018 (ENE) en la región $r$
$\lambda$	: Porcentaje de selección asignado para cada región (70,0% o 85,0%)
$SE(\hat{p}_{r0})$	: Error estándar de la estimación de la proporción de microemprendedores por cuenta propia en la región $r$ . Corresponde a la raíz cuadrada de la varianza de la estimación
$Z_{1-\alpha/2}$	: Percentil de nivel $(1 - \alpha/2)$ de la distribución Normal, correspondiente a una estimación intervalar de $(1 - \alpha)$ de confianza

#### **Etapa 1:**

Utilizando la base de datos del levantamiento de la ENE del trimestre marzo, abril y mayo (MAM) 2018, se obtuvieron las estimaciones a nivel regional de: parámetro de interés regional ( $\hat{p}_{r0}$ ), error estándar ( $SE(\hat{p}_{r0})$ ) y total de unidades logradas ( $m_{r0}$ ).

#### **Etapa 2:**

Se calcularon los errores absolutos regionales ( $ea_{r0}$ ) a partir de los errores estándar obtenidos en MAM 2018 ( $SE(\hat{p}_r)$ ) estableciendo un nivel de confianza de 95% bajo una distribución normal ( $Z_{1-\alpha/2}$ ). El cálculo se observa en la ecuación (1):

$$ea_{r0} = 1,96 \cdot SE(\hat{p}_{r0}) \quad (1)$$

#### **Etapa 3**

El tamaño muestral de viviendas a nivel regional ( $m_{r1}$ ) fue equivalente a 70,0% de las viviendas disponibles en el marco de muestreo, exceptuando las regiones de: Antofagasta,

Ñuble y Magallanes, donde el tamaño corresponde a 85,0% de lo disponible, producto de la alta variabilidad de la variable de interés. El cálculo del tamaño se observa en la ecuación (2):

$$m_{r1} = m_{r0} \cdot \lambda^{14} \quad (2)$$

#### **Etapa 4**

Una vez obtenidos los tamaños regionales, se aplicó un factor que corresponde a la tasa de no respuesta (*tnr*) obtenida en la V EME, con la finalidad de salvaguardar la precisión de la estimación del parámetro de interés, debido a la posibilidad de que no se lograra el total de unidades a encuestar, por diversas razones, tales como: rechazos, moradores ausentes, edificaciones que no forman parte de la población objetivo, entre otras.

Para esta versión de la encuesta, se utilizó una tasa de no respuesta de 15,0% a nivel regional (según tasas obtenidas en la V EME) salvo para la Región Metropolitana, donde se contempló 20,0%<sup>15</sup>. Luego, se obtuvo el tamaño muestral con sobremuestreo, o ajustado por no respuesta, para cada región como se precisa en la ecuación (3).

$$m_{r2} = \frac{m_{r1}}{1 - tnr} \quad (3)$$

La siguiente tabla presenta los errores esperados, asociados a la estimación del parámetro de interés obtenido en la V EME, para un tamaño muestral objetivo de 7.086 viviendas<sup>16</sup> a nivel nacional.

---

<sup>14</sup> Porcentaje de selección asignado, según se haya definido para cada región.

<sup>15</sup> Debido a que en el levantamiento de las primeras 2 Submuestras se observaron niveles de no respuesta mayores al 20% en la Región Metropolitana, en la tercera submuestra se optó por seleccionar todas las viviendas disponibles para aumentar el número de encuestas logradas y acercarse más a la muestra objetivo. Al realizar este procedimiento se aumentaron en 12 las viviendas seleccionadas en esta región, desde 2.024 a 2.036.

<sup>16</sup> La estimación de los errores se obtiene a partir del tamaño objetivo y no del tamaño con sobremuestreo, pues dadas las pérdidas naturales de unidades muestrales, se espera obtener un tamaño de 7.086 viviendas.



**Tabla 6: Errores esperados asociados al parámetro de interés según tamaño muestral objetivo**

Nivel	Estimación V EME	Error absoluto propuesto VI EME	Error relativo propuesto VI EME	Tamaño objetivo VI EME	Tamaño con sobremuestreo VI EME
Nacional	86,0%	1,3%	1,5%	7.086	8.469 <sup>17</sup>
Arica y Parinacota	85,1%	5,2%	6,1%	274	322
Tarapacá	93,7%	3,3%	3,5%	248	292
Antofagasta	81,3%	7,6%	9,4%	212	249
Atacama	87,1%	5,9%	6,8%	206	242
Coquimbo	87,3%	4,4%	5,0%	494	581
Valparaíso	82,1%	3,4%	4,1%	928	1.092
Metropolitana	86,9%	2,4%	2,8%	1.619	2.036
O'Higgins	82,1%	5,4%	6,5%	360	424
Maule	85,0%	4,7%	5,6%	430	506
Ñuble	87,4%	5,9%	6,7%	155	182
Biobío	88,4%	3,3%	3,7%	644	758
La Araucanía	87,7%	4,5%	5,2%	468	551
Los Ríos	86,3%	5,6%	6,5%	236	278
Los Lagos	83,9%	4,2%	5,0%	519	611
Aysén	83,7%	5,7%	6,9%	188	221
Magallanes	74,7%	8,2%	11,0%	105	124

Fuente: INE, VI Encuesta de Microemprendimiento, 2019

La muestra efectiva de la VI EME fue construida utilizando el proceso descrito más arriba, pero utilizando como marco muestral los datos entregados por el levantamiento del trimestre marzo, abril y mayo de 2019 de la EME. A partir de este punto, los datos de muestra corresponden al muestreo efectivo.

<sup>17</sup> El aumento en el número de viviendas seleccionadas en la región ocasionó un aumento en la selección total a nivel nacional de 8.457 a 8.469.

#### 4.3.3. Distribución de la muestra por submuestras

Como la muestra ENE estuvo subdividida en tres meses o períodos de levantamiento, el mes de referencia de marzo se levantó en mayo en la EME, el de abril en junio y el mayo en Julio, esto permitió evitar que pasara mucho tiempo entre el levantamiento de la ENE y de la EME para así tener una menor atrición<sup>18</sup> de la muestra.

Una vez obtenido el tamaño muestral nacional y regional, la muestra de la VI EME, se distribuyó en tres partes iguales, o lo más similares posibles, en los siguientes meses de levantamiento: mayo, junio, julio. La tabla 7 muestra la distribución de la muestra según mes de levantamiento y región.

**Tabla 7: Total de viviendas seleccionadas según región y mes de levantamiento**

Región	Mes de levantamiento VI EME			Viviendas seleccionadas
	Mayo	Junio	Julio	
<b>Nacional</b>	<b>2.819</b>	<b>2.819</b>	<b>2.831</b>	<b>8.469</b>
<b>Arica y Parinacota</b>	107	108	107	322
<b>Tarapacá</b>	97	99	96	292
<b>Antofagasta</b>	84	80	85	249
<b>Atacama</b>	80	82	80	242
<b>Coquimbo</b>	193	194	194	581
<b>Valparaíso</b>	364	364	364	1.092
<b>Metropolitana</b>	675	674	687	2.036
<b>O'Higgins</b>	142	141	141	424
<b>Maule</b>	168	169	169	506
<b>Ñuble</b>	61	60	61	182
<b>Biobío</b>	253	253	252	758
<b>La Araucanía</b>	183	179	189	551
<b>Los Ríos</b>	93	93	92	278
<b>Los Lagos</b>	204	205	202	611
<b>Aysén</b>	73	75	73	221
<b>Magallanes</b>	42	43	39	124

Fuente: INE, VI Encuesta de Microemprendimiento, 2019

La tabla 8 presenta los errores observados, a nivel nacional y regional, asociados a la prevalencia del parámetro de interés, para un tamaño muestral efectivo de 7.301 viviendas, según levantamiento de la VI EME.

<sup>18</sup> Acumulación de pérdida de información por la no respuesta que se presenta en estudios sobre unidades en el tiempo por parte de los participantes (Jones, Koolman, & Rice, 2006).

**Tabla 8: Errores observados asociados al parámetro de interés según número de viviendas que responden<sup>19</sup>**

Región	Estimación VI EME	Tamaño efectivo	Error absoluto	Error relativo
Nacional	85,5%	7.301	1,2%	1,4%
Arica y Parinacota	85,6%	278	5,6%	6,5%
Tarapacá	88,3%	281	5,3%	5,9%
Antofagasta	85,3%	233	9,1%	10,7%
Atacama	87,3%	227	4,8%	5,5%
Coquimbo	85,3%	538	3,6%	4,3%
Valparaíso	86,5%	946	3,0%	3,5%
Metropolitana	84,9%	1.583	2,3%	2,7%
O'Higgins	88,8%	368	4,3%	4,8%
Maule	88,3%	458	2,8%	3,2%
Ñuble	81,0%	154	4,5%	5,6%
Biobío	81,3%	662	5,9%	7,2%
La Araucanía	89,6%	483	2,7%	3,0%
Los Ríos	87,7%	253	5,6%	6,4%
Los Lagos	82,7%	535	2,7%	3,2%
Aysén	80,9%	196	6,0%	7,5%
Magallanes	76,3%	106	7,1%	9,3%

Fuente: INE, VI Encuesta de Microemprendimiento, 2019

#### 4.4. Selección de unidades muestrales

La selección de unidades muestrales se realizó en 2 etapas. Primero sobre las viviendas que contienen al menos un microemprendedor y luego, al interior se seleccionan los microemprendedores.

##### 4.4.1. Selección de viviendas

La selección de viviendas se hizo de forma sistemática con igual probabilidad de selección, para todas las viviendas al interior de cada región.

La selección fue implementada en el *software* estadístico SPSS, en el módulo de análisis “Muestras Complejas” específicamente, en el procedimiento “seleccionar una muestra”, fijando una semilla aleatoria, a fin de que pueda ser replicable en cualquier momento.

Sean:

$m_r$  : Número de viviendas a seleccionar en la región  $r$ .

$M_r$  : Número de viviendas que contiene la región  $r$ .

Para la selección de  $m_r$  viviendas, el *software* ejecuta los siguientes pasos:

<sup>19</sup> La estimación de los errores se obtiene a partir del tamaño efectivo, utilizando como variable de estratificación pseudo estratos (Varstrat) y variable de conglomeración pseudo conglomerados (Varunit). El detalle de la construcción de los pseudo estratos y pseudo conglomerados se encuentra en el Informe metodológico del diseño muestral.

### **Paso 1**

En primera instancia, se ordenaron en forma ascendente todas las viviendas según región, estrato, área, comuna, sección, distrito censal, zona censal, número de manzana y número de orden de vivienda dentro de la manzana.

### **Paso 2**

Al interior de cada región se calculó el período ( $k$ ) que corresponde a:

$$k = M_r/m_r.$$

Notar que " $k$ " puede ser un número real, no entero (puede tener decimales).

### **Paso 3**

Luego se determinó el arranque " $A$ " o primera selección, que corresponde a una semilla aleatoria propia para la encuesta.

### **Paso 4**

Posteriormente se sumó sucesivamente el período " $k$ " al arranque " $A$ " para obtener distintos valores, los que dieron origen a la selección de unidades de la siguiente forma: " $A$ ", " $A + k$ ", " $A + 2k$ ", " $A + 3k$ ", ..., " $A + (m_r - 1)k$ ".

La primera vivienda seleccionada fue " $A$ " y es un número entero, la segunda fue el redondeo de " $A + k$ ", la tercera fue el redondeo de " $A + 2k$ " y así sucesivamente, hasta la  $m_r$  selección, dada por el redondeo de " $A + (m_r - 1)k$ ".

Luego, la probabilidad de inclusión de la  $j$  – ésima vivienda dentro de la  $r$  – ésima región, fue igual a:

$$P_r(j|r) = \frac{m_r}{M_r} \quad (4)$$

#### 4.4.2. Selección de informantes

Una vez seleccionadas las viviendas, al interior de sus hogares se identificó a los microemprendedores y las actividades económicas en que estos se desenvolvían. Luego, se seleccionaron de forma aleatoria y con igual probabilidad, tantos microemprendedores como actividades distintas identificadas dentro del hogar, es decir, en caso de encontrar más de un microemprendedor en el hogar ejecutando la misma actividad económica, se debía seleccionar sólo a un representante por actividad dentro del hogar.

## 5. METODOLOGÍA DE RECOLECCIÓN DE DATOS

El presente capítulo detalla las metodologías implementadas en el proceso de recolección de datos. El primer subapartado se detallan técnicas, medios y las fechas de la recolección oficial. En el segundo se describe el proceso de diseño del instrumento de recolección oficial por medio del levantamiento de una prueba piloto. En el tercero, se exponen las metodologías utilizadas para evaluar el instrumento de recolección piloteado. El cuarto subapartado detalla el método de capacitación, de encuestadores, supervisores y analistas; y finalmente en el quinto subapartado, se detallan las metodologías de supervisión utilizadas para resguardar la calidad de la información recopilada.

### 5.1. Técnica y Métodos de recolección de datos

#### 5.1.1. Principales técnicas de recolección de datos.

El objetivo principal de la VI Encuesta de Microemprendimiento es caracterizar profundamente a los microemprendimientos que se desarrollan a nivel nacional ya sean formales o informales. Para poder estudiar a toda la población microempresaria se optó por realizar una encuesta en hogares lo que permite observar emprendimientos que se desarrollan al margen de la formalidad. Esta técnica de recolección de datos permite obtener información de las unidades económicas de menor tamaño con mayor facilidad y precisión, pues los micronegocios o microemprendimientos muchas veces no se encuentran en los listados oficiales como los del Servicio de Impuestos Internos.

#### 5.1.2. Medio de recolección de información utilizado.

En esta oportunidad, tal como en las versiones anteriores, la VI EME ocupó el método de captura PAPI (paper and pencil interview), utilizándose un formulario de 24 páginas que incluyó la hoja de ruta y sus observaciones, el cuestionario propiamente tal y las observaciones de terreno.

Se realizaron entrevistas presenciales, a los microempresarios seleccionados a partir del marco muestral que nació de la ENE de marzo, abril y mayo de 2019. Si bien la Encuesta Nacional de Empleo puede ser contestada por un informante idóneo que entrega información para todos quienes habitan en una vivienda, para la EME se encuestó directamente al microempresario identificado por la ENE, debido a la especificidad de las preguntas.

#### 5.1.3. Período de referencia

El período de referencia de donde se obtiene la muestra es el trimestre marzo, abril, mayo de 2019 de la Encuesta Nacional de Empleo. El trabajo de campo de la EME se realizó durante el trimestre mayo, junio, julio de 2019.

Adicionalmente, el cuestionario de la VI EME cuenta con preguntas que piden al informante que se sitúe en diferentes períodos de referencia siendo dos horizontes temporales los más frecuentes en el cuestionario: los últimos doce meses (utilizado principalmente en las preguntas de gastos, ingresos y ganancias) y en la actualidad, es decir, al momento de contestar la encuesta. Otros períodos de referencia presentes en la encuesta son el último

mes trabajado, el mes anterior, los últimos dos años y todo el tiempo que ha trabajado como microempresario.

#### 5.1.4. Fechas de recolección de datos y procesos complementarios

El proceso de recolección de datos se inició con la capacitación de los encuestadores y supervisores el 23 de abril. Inmediatamente terminada la capacitación comenzó la aplicación de la primera de las tres submuestras que conformaron la muestra de la VI EME.

A pocos días de que se diera el vamos al levantamiento de información, el equipo técnico a cargo de la EME inició un ciclo de visitas técnicas con el objetivo de supervisar y observar el trabajo que realizaban encuestadores y supervisores. Estas visitas se extendieron por casi la totalidad de mayo y junio de 2019 y a lo largo del país.

En paralelo, equipo técnico también realizó capacitaciones para el equipo de analistas de a cargo de recepcionar las encuestas y llevar a cabo el análisis de validaciones de primer y segundo orden, desarrollando un primer procesamiento y edición de la información previo a la revisión del equipo técnico.

**Diagrama 1: Línea de tiempo de trabajo de recolección de datos.**



Fuente: INE, VI Encuesta de Microempresamiento, 2019

Es importante señalar que las fases del trabajo de recolección y posterior procesamiento de la VI EME mencionadas en los párrafos anteriores, estuvieron interrelacionadas con las otras, de manera tal que, si en la realización de una de estas tareas se detectaron errores, estos fueron corregidos, llegando incluso a devolverse formularios a los equipos a cargo del levantamiento.

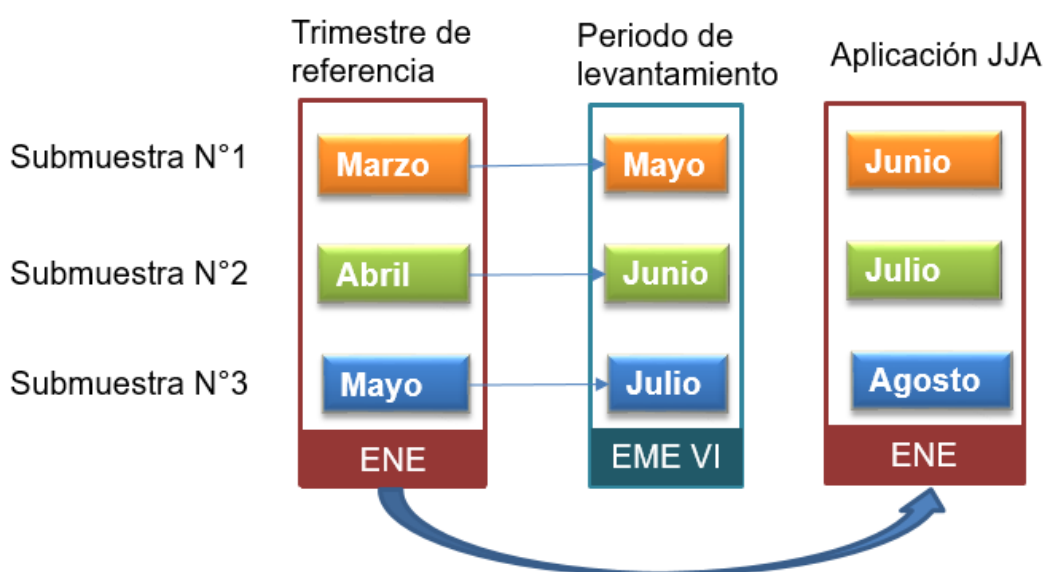
La muestra de la VI EME estuvo dividida en tres submuestras, cada una de ellas encuestó a un grupo específico de informantes. Esta subdivisión del trabajo de recolección fue realizada para no interferir el trabajo de recolección de la ENE. La Encuesta de Microempresamiento es una segunda fase a una submuestra de la Encuesta Nacional de Empleo, la cual corresponde a los ocupados clasificados como empleadores o trabajadores independientes.

La ENE posee un diseño muestral en el que se entrevista una vivienda en varias ocasiones, número que varía dependiendo del estrato en el que se encuentra la vivienda (urbano, RAU o rural), pudiendo ser entrevistada 6, 9 o hasta 12 veces. Cada aplicación está separada

por tres meses de la anterior, lo que implica que una vivienda forma parte de la muestra de la ENE durante 18, 27 o 36 meses.

En este contexto la mayor parte de la EME se aplica sobre muestra activa de la ENE (5/ 6 de la muestra urbana y un porcentaje mayor de la muestra RAU y Rural serán vueltos a responder), por lo que la recolección de datos debe cuidar que este no interrumpa el ciclo de visitas realizadas por la ENE. Esta es la razón por la que quienes fueron originalmente entrevistados durante marzo de 2019 en la ENE, formaron parte de la primera submuestra de la EME realizada durante mayo de 2019, previo a que fuesen nuevamente contactados por la ENE en junio de 2019. La siguiente figura resume en qué momento se encuestó a cada integrante de la muestra dependiendo del mes que fue entrevistado por la ENE.

**Diagrama 2: Esquema de participación de los informantes en la ENE y EME.**



Fuente: INE, VI Encuesta de Microemprendimiento, 2019

5.1.5. Identificar el tipo de entrevista y la confirmación del equipo de trabajo (encuestadores, supervisores y proporción por región)

A lo largo del país fueron 52 encuestadores los que se encargaron de recolectar la información que forma parte de la VI EME, a ellos se sumaron 19 supervisores que monitorearon el avance del trabajo y la calidad de este. A continuación, en la Tabla 9, se expone la distribución geográfica de los encuestadores y supervisores que tomaron parte de esta versión de la Encuesta de Microemprendimiento.

**Tabla 9: Encuestadores y supervisores por región**

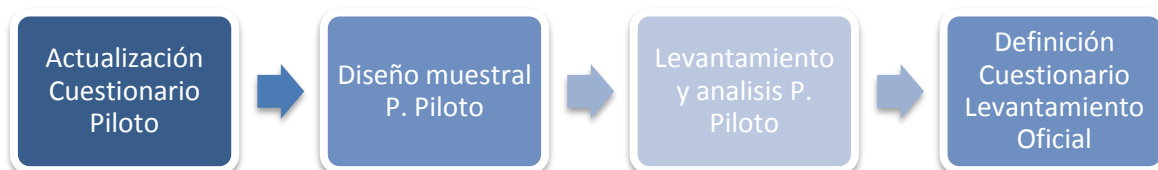
Región	Encuestadores	Supervisores
Arica y Parinacota	2	1
Tarapacá	2	1
Antofagasta	2	1
Atacama	2	1
Coquimbo	4	1
Valparaíso	6	2
Metropolitana	11	3
O'Higgins	3	1
Maule	4	1
Ñuble y Biobío	5	2
La Araucanía	3	1
Los Ríos	2	1
Los Lagos	4	1
Aysén	1	1
Magallanes	1	1
<b>Nacional</b>	<b>52</b>	<b>19</b>

Fuente: INE, VI Encuesta de Microemprendimiento, 2019

## 5.2. Diseño del instrumento de captura de información

El proceso de generación del instrumento de captura de datos contó con la aplicación de una prueba piloto para testear el cuestionario diseñado. Este cuestionario consideró algunas nuevas áreas de interés definidas entre el Ministerio de Economía, instituciones interesadas y el Instituto Nacional de Estadísticas (INE). El proceso de diseño del nuevo instrumento a utilizar en la VI EME se puede resumir en cuatro etapas:

**Diagrama 3: Etapa de diseño de materiales**



Fuente: INE, VI Encuesta de Microemprendimiento, 2019



### 5.2.1. Actualización cuestionario Piloto

A partir de la demanda de información realizada por usuarios externos al Ministerio de Economía, Fomento y Turismo y el INE, para esta sexta versión de la Encuesta de Microemprendimiento se incorporaron nuevas áreas de interés al cuestionario, generándose modificaciones que fueron testeadas en una prueba piloto aplicada a una muestra reducida de informantes.

Para ello se diseñó un nuevo cuestionario (ver tabla 10) el cual conservó en gran parte la estructura y el contenido de la versión anterior de la encuesta (V EME 2017), incorporando dos nuevos módulos, uno de ellos correspondiente al Módulo D de “Gastos, Ingresos y Ganancias del Negocio” el cual proviene de una reformulación de las preguntas de ingresos y gastos pertenecientes al Módulo D “Características del Negocio” de la V EME. Además, se incorporó un último Módulo L “Trabajo No Remunerado” que tiene por finalidad estimar el tiempo dedicado a recoger información sobre actividades de trabajo doméstico y cuidados que los informantes realizan sin recibir una remuneración.

Debido a la incorporación de nuevas áreas de interés y el cambio de metodología de captura en ciertas variables, el número de preguntas aumentó de 89 a 120<sup>20</sup>, incluyendo 52 preguntas nuevas y eliminando 21 de la versión anterior<sup>21</sup>. Este trabajo de actualización se llevó a cabo durante los meses de octubre y diciembre de 2018 y se sustentó en reuniones sostenidas entre la Subsecretaría de economía y empresas de menor tamaño (en adelante Subsecretaría) y usuarios externos, seminarios de difusión de resultados y las discusiones técnicas llevadas a cabo entre la Subsecretaría y el INE en base a los marcos conceptuales vigentes propuesto por OIT establecidos principalmente en la décimo novena resolución de la CIET y el nuevo clasificador de Actividades de uso del tiempo para América Latina (CAUTAL).

---

<sup>21</sup> Es importante considerar que gran parte de las preguntas eliminadas corresponden a preguntas de ingresos y gastos pertenecientes al módulo D (11 preguntas), las cuales sirvieron como referencia para crear el nuevo módulo D “Gastos, Ingresos y Ganancias del negocio”. Por otra parte, un grupo importante de las preguntas “eliminadas” en realidad fueron localizadas en nuevos módulos, en particular de las 7 preguntas nuevas del módulo B, 6 provienen de otros módulos (4 del Módulo A y 2 del Módulo D).

**Tabla 10: Cambios cuestionario Prueba Piloto VI EME**

AREAS DE ESTUDIO	MÓDULO		N° PREGUNTAS			
	V EME	VI EME	V EME	VI EME (Piloto)	Eliminadas	Agregadas
Situación Laboral	A	A	6	12	4	10
Historia Laboral	B	B	7	14		7
Capacitación	C	J	5	5		
Características de la empresa	D	C	19	8	11	
Gastos, Ingresos y Ganancias		D		16		16
Situación de Formalidad	E	E	10	13		1
Generación de empleo	F	F	8 (15)	8 (17)		(2)
Relación con el sistema Financiero	G	G	12	14	6	8
Activos	H	H	4	5		1
Tecnología e información	I	I	3	3		
Entorno del negocio	J	K	15	17	0	2
Trabajo No Remunerado		L		5		5
<b>Total</b>			<b>89 (96)</b>	<b>120 (129)</b>	<b>21</b>	<b>50 (2)</b>

Fuente: INE, VI Encuesta de Microemprendimiento, 2019

El aumento en la cantidad total de preguntas del cuestionario respondió a una necesidad de profundizar en la caracterización de los microemprendedores, especialmente en las preguntas que permiten construir las variables de gastos, ingresos y ganancias del negocio, la motivación emprendedora, el trabajo no remunerado y algunas preguntas de caracterización de los microemprendedores que cambian de situación ocupacional.

Principalmente los cambios diseñados tuvieron relación con cuatro áreas de indagación. La primera de ellas tiene relación con el sistema financiero (Módulo G) donde se levantó la necesidad de contar con más información sobre la solicitud de crédito o préstamos independiente de la institución.

Por otra parte, se diseñaron una serie de cambios en las preguntas de ingresos y gastos con el objetivo de mejorar la captura de los ingresos y gastos del negocio, incorporándose diferentes metodologías de indagación para los ingresos provenientes del sector agrícola y no agrícolas. Además, se incluyeron preguntas respecto a los ingresos derivados del arriendo de equipos, vehículos o herramientas del negocio e ingresos de actividades complementarias (pensión, subsidios, seguros, etc.).

Producto del aumento en la ocupación de las mujeres, su importancia dentro de los microemprendimientos y las grandes brechas en el tiempo dedicado al trabajo no remunerado, para esta versión se incluyeron algunas preguntas que tienen el objetivo de estimar el tiempo dedicado al trabajo no remunerado, visibilizando las brechas en este tipo de trabajo por sexo.

Finalmente, en la Prueba Piloto se testearon nuevas preguntas que buscan mejorar la caracterización de los informantes que cambian de situación ocupacional (dejaron de ser trabajadores independientes entre la aplicación de la ENE y la EME), indagando en los motivos de esta decisión y su actual condición ocupacional.

#### 5.2.2. Diseño prueba piloto

Al igual que en el levantamiento oficial, el diseño muestral de la Prueba Piloto en la VI Encuesta de Microemprendimiento 2019 correspondió a un diseño probabilístico y bifásico, donde la primera fase fue el levantamiento de la Encuesta Nacional de Empleo (ENE) y la segunda una Encuesta mixta a trabajadores independientes dueños de una Microempresa.

#### I. Características generales del diseño muestral

##### a) Objetivo diseño muestral

El objetivo principal del diseño muestral de la Prueba Piloto fue analizar el instrumento de medición en un grupo de unidades de observación, que permitan estudiar la comprensión ante distintos perfiles de informantes. Particularmente para esta versión fue de especial interés testear las preguntas del Módulo de ingresos a microemprendedores del sector agrícola.

##### b) Población objetivo

La población objetivo de la Prueba Piloto la conformaron todas las personas de 15 años o más, residentes habituales de las viviendas particulares ocupadas, que fueron clasificadas como Microemprendedores<sup>22</sup> en la Encuesta Nacional de Empleo, en el período septiembre, octubre y noviembre de 2018 (SON 2018) y junio, julio y agosto de 2018 (JJA 2018).

#### II. Características del Marco Muestral

##### a) Cobertura del marco muestral

La muestra de la Prueba Piloto cubrió el área urbana y rural de la Región Metropolitana y la Región de O'Higgins. Esta es la primera vez que la muestra incluyó viviendas de regiones distintas a la Metropolitana. Esta decisión tuvo el fin de aumentar la probabilidad de seleccionar microemprendedores dedicados a actividades agrícolas, lo que facilitaría el testeado de las nuevas preguntas incluidas en el cuestionario piloto que permiten capturar las particularidades de los ingresos provenientes de este sector.

##### b) Estratificación del marco muestral

Al ser una encuesta en segunda fase, la estratificación del marco siguió las mismas especificidades de la Encuesta Nacional de Empleo, es decir, consideró tres tipos de

---

<sup>22</sup> Trabajador por cuenta propia o empleador con hasta 10 trabajadores (incluyéndose).

estratos Ciudad o grandes centros urbanos (CD), Resto de Área Urbana (RAU), Área Rural (R).

### III. Cálculo y tamaño de la muestra

La prueba piloto tuvo por objetivo testear en terreno el desempeño del instrumento de medición, con el fin de revisar el nivel de entendimiento y comprensión que tiene el entrevistado. Por lo tanto, el tamaño muestral definido no consideró estimaciones confiables de ciertas variables de interés, sino más bien, que la cantidad de individuos que respondiesen la encuesta fuera la adecuada para realizar análisis descriptivos

En este contexto y, teniendo en consideración otro tipo de restricciones como el tiempo disponible para levantar la prueba piloto y los costos operativos asociados a éste, se definió que el tamaño muestral objetivo fuese de 150 viviendas las cuales permitían aplicar la encuesta a distintos perfiles de informantes. Considerando una tasa de no respuesta de 41% proveniente de la encuesta piloto anterior, se proyectó en un tamaño de 254 viviendas con sobremuestreo para alcanzar el número deseado de entrevistas. El objetivo del sobremuestreo fue minimizar potenciales problemas de participación de alguno de los perfiles necesarios a analizar ya sea porque las viviendas en las que residen se encontraban no logradas por motivos como el de moradores ausentes, cambio de estado, rechazos, etc.

#### a) Selección de las unidades muestrales

La selección de unidades muestrales se realizó en dos etapas. Primero sobre las viviendas que contienen al menos un microemprendedor y segundo, al interior se seleccionan los microemprendedores.

#### b) Selección de viviendas

La selección de las viviendas donde residen los microemprendedores se realizó sobre todas aquellas que cumplieron su última visita en la muestra de la ENE en el trimestre móvil septiembre, octubre y noviembre (SON) de 2018. Como, a partir de este criterio, no fue posible lograr el tamaño necesario de viviendas en las que residían microemprendedores agrícolas, se realizó una selección de unidades agrícolas<sup>23</sup> que cumplieron la penúltima visita en dicho trimestre y, una selección de aquellas viviendas que cumplieron la última visita en el trimestre móvil julio, junio, agosto (JJA) de 2018. Además, dada la baja participación de viviendas con microemprendedores agrícolas en la RM, se incluyeron viviendas de la región de O'Higgins con microemprendedores que se dedicaban a este tipo de actividades. En detalle, se lograron seleccionar 46 viviendas<sup>24</sup> con al menos 1 microemprendedor agrícola.

---

<sup>23</sup> Las actividades agrícolas se obtienen de la categoría "1" de la variable B14 de la encuesta, cuya clasificación corresponde a: agricultura, caza y pesca.

<sup>24</sup> Lo que corresponde a 18% del total de la muestra.

**Tabla 11: Distribución de muestra agrícola por región y trimestre de referencia**

Tipo muestra	Región		Total
	O'Higgins	Metropolitana	
Última visita JJA	8	4	12
Última visita SON	10	6	16
Penúltima visita SON	10	8	18
<b>Total</b>	<b>28</b>	<b>18</b>	<b>46</b>

Fuente: INE, VI Encuesta de Microemprendimiento, 2019

Finalmente, en el siguiente cuadro se observa la distribución de la muestra final para el levantamiento de la Prueba Piloto de la VI EME, según tipo de estrato y región.

**Tabla 12: Prueba Piloto. Distribución de viviendas muestra final según región y trimestre de referencia**

Trimestre	Región		Total
	O'Higgins	Metropolitana	
JJA	8	4	12
SON	20	222	242
<b>Total</b>	<b>28</b>	<b>226</b>	<b>254</b>

Fuente: INE, VI Encuesta de Microemprendimiento, 2019

### c) Selección de informantes

Una vez seleccionadas las viviendas, al interior de los hogares se eligieron un número de informantes igual al número de las ramas de actividad presentes en el hogar. La mayoría de los informantes agrícolas se obtuvieron de la región de O'Higgins y de la muestra adicional que cumplió su última visita en julio - agosto y penúltima en septiembre- noviembre de la región Metropolitana. La siguiente tabla muestra la distribución de informantes seleccionados por región y sector de actividad económica

**Tabla 13: Distribución de muestra según región y sector de actividad**

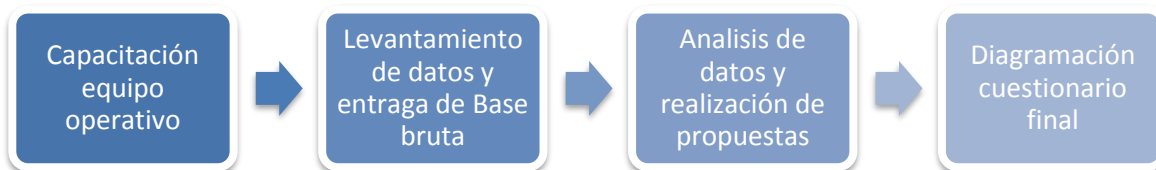
Sector	Región		Total
	O'Higgins	Metropolitana	
No Agrícola	5	258	263
Agrícola	31	19	50
<b>Total</b>	<b>36</b>	<b>277</b>	<b>313</b>

Fuente: INE, VI Encuesta de Microemprendimiento, 2019

### 5.2.3. Levantamiento y análisis prueba piloto

La fase del levantamiento y análisis de la prueba piloto fue realizada entre los meses de febrero y marzo del 2019 y consideró principalmente tres actividades: Capacitación equipo operativo, levantamiento y procesamiento de datos, análisis de datos y realización de propuestas. A estas tres, se sumaron las labores de diagramación del cuestionario final.

**Diagrama 4: Proceso construcción cuestionario final**



Fuente: INE, VI Encuesta de Microemprendimiento, 2019

El levantamiento de datos para el piloto de la VI Encuesta de Microemprendimiento, tal como ya se mencionó, contempló la realización de una encuesta a los informantes previamente seleccionados en las regiones Metropolitana y de O'Higgins.

El trabajo de campo fue desarrollado entre los días 4 y 24 de febrero de 2019, período que se dividió en dos semanas para el levantamiento de la información y una para la recuperación. Las 28 viviendas con microempresarios agrícolas de la de O'Higgins fueron de levantadas por dos encuestadores entre los días 18 y 20 de febrero según lo planificado.

Respecto a la organización del trabajo operativo del levantamiento piloto esta consideró las mismas etapas que el levantamiento oficial: Levantamiento de datos, supervisión digitación y análisis.

En cuanto al procesamiento de los datos del levantamiento piloto, por restricciones temporales sólo se realizaron validaciones de primer nivel, las que hacen referencia a la detección de errores de rango y omisión de información. Una vez realizada todas las actividades detalladas para cada unidad incluida en la muestra, se dio por cerrado el trabajo operativo, con la entrega de la base de datos final del levantamiento piloto.

Tras la finalización del trabajo de campo y la recepción de la base de datos con el total de casos del piloto, se realizaron algunas reuniones en las que se definieron los contenidos del cuestionario con el Ministerio de Economía y con el equipo técnico de la Encuesta de Uso del tiempo (ENUT), quienes realizaron recomendaciones al módulo de trabajo no remunerado. El detalle de la definición del cuestionario oficial se encuentra en el siguiente apartado:

### 5.2.4. Definición cuestionario oficial

Uno de los hallazgos o problemas más importantes del cuestionario piloteado fue el considerable aumento en su extensión, aumentando de 89 a 120 el número de preguntas respecto de su versión anterior. Si bien parte de las preguntas incluidas fueron aplicadas

solamente a quienes cambiaban de condición de actividad (A1=3), la duración promedio entre quienes mantenían su condición de trabajador independiente aumentó de 30 a 44 minutos:

**Tabla 14: Duración de entrevistas según Formato**

Duración	V EME	VI EME*
Promedio	30,8 minutos	49 minutos
Máximo	90 minutos	100 minutos
Mínimo	1 minuto	22 minuto*

\*Excluye casos que cambian de situación ocupacional (A1=3)

Fuente: INE, VI Encuesta de Microemprendimiento, 2019

El aumento en la duración del cuestionario fue provocado principalmente por la longitud de los módulos de D “Gastos, Ingresos y Ganancias”, G “Relación con el sistema financiero” y L “Trabajo No Remunerado”. Por ello, al momento de construir el cuestionario final se decidió realizar un análisis de la comprensión y uso de las preguntas del cuestionario, con el objetivo de evaluar la reducción en el número de estas y con ello disminuir el tiempo de aplicación.

Esta evaluación implicó la eliminación de 27 preguntas del cuestionario. La tabla 15 detalla el total de preguntas del cuestionario de la V y VI EME, tanto del piloto como del levantamiento oficial, incluyendo también el detalle de las preguntas agregadas y eliminadas por módulo:

**Tabla 15: cambios cuestionario Prueba Piloto VI EME**

AREAS DE ESTUDIO	V EME	VI EME (Piloto)	VI EME (L. Oficial)	Eliminadas (en relación con Piloto)	Agregadas (en relación con Piloto)
Situación Laboral	6	12	12	0	0
Historia Laboral	7	14	9	5*	0
Características de la empresa	19	8	9	0	1**
Gastos, Ingresos y Ganancias		16	12	5**	1
Situación de Formalidad	10	13	10	3*	0
Generación de empleo	8	8	4	4	0
Relación con el sistema Financiero	12	14	13	1	0
Activos	4	5	4	1	0
Tecnología e información	3	3	4	0	1
Capacitación	5	5	4	1	0
Entorno del negocio	15	17	13	4	0
Trabajo No Remunerado		5	3	3	1
<b>Total</b>	<b>89</b>	<b>120</b>	<b>97</b>	<b>27</b>	<b>4</b>

\* De las preguntas eliminadas en los módulos B y E, 3 de ellas en realidad fueron fusionadas con otras preguntas de estos módulos por lo que la información recaba por éstas no se pierde.

\*\* La pregunta agregada en el Módulo C, realmente no corresponde a la incorporación de una pregunta al cuestionario sino al traspaso de una pregunta del módulo D al C.

Fuente: INE, VI Encuesta de Microemprendimiento, 2019

### 5.3. Prueba y evaluación del instrumento

En esta sección se presentan las distintas metodologías que fueron utilizadas para evaluar el instrumento en terreno junto con algunos resultados. Entre ellas se encuentran tres metodologías cualitativas y una metodología cuantitativa, las cuales en conjunto permitieron analizar los diversos hallazgos o problemas desde la óptica de los distintos participantes de la encuesta.

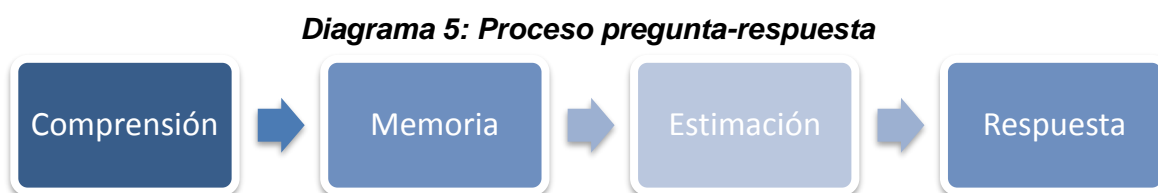
#### 5.3.1. Entrevistas cognitivas

La primera técnica que se usó para analizar el cuestionario de la prueba piloto fue la entrevista cognitiva, metodología que busca obtener información relevante para el perfeccionamiento del cuestionario y posibilitar la generación de mejoras para la posterior implementación oficial.

##### I. Metodología entrevistas cognitivas

Esta técnica se basa en la psicología cognitiva, específicamente en la teoría que plantea que el proceso de respuesta que realizan las personas presenta cuatro distintas etapas: comprensión, memoria, estimación y respuesta. Estas etapas no son siempre lineales o consecutivas, pero cada una de ellas da cuenta de un proceso cognitivo particular donde podrían presentarse problemas con algunas preguntas del cuestionario.

A nivel general se podría describir el flujo del proceso pregunta-respuesta en base al siguiente diagrama:



Fuente: Cognitive Interviewing, Gordon Willis (2004)

El primer tipo de hallazgo corresponde al de comprensión, el cual se manifiesta cuando un informante presenta problemas de entendimiento del enunciado o de las alternativas de respuesta por parte del informante. Generalmente este hallazgo se observa cuando el encuestador debe repetir la pregunta o explicaren detalle el sentido de esta.

Un segundo problema en el proceso pregunta-respuesta es el relativo a la memoria, es decir, se presenta en informantes que evidencian ciertas dificultades al momento de responder la pregunta pues deben de recordar alguna fecha o evento que transcurrieron hace un tiempo.



El tercer problema ocurre cuando un informante presenta dificultades al momento de realizar algún cálculo o estimación para responder la pregunta del encuestador. Por ejemplo, en el caso de la pregunta de antigüedad laboral (¿cuántos años lleva trabajando en esa actividad o negocio?).

El cuarto problema se presenta en la última etapa del proceso preguntas y respuestas, en informantes que presentan ciertas dificultades al entregar una respuesta al encuestador. Esto puede ser ocasionado porque las alternativas presentes en el cuestionario no dan cabida a la respuesta del informante o bien porque el informante teme revelar la respuesta real (por ejemplo, en preguntas de ingresos o situación de formalidad).

Las entrevistas cognitivas se realizaron siguiendo una pauta de entrevista semi-estructurada<sup>25</sup>, con preguntas del cuestionario y sondeos cognitivos respecto a las mismas. Se realizaron ocho entrevistas cognitivas para esta versión de la EME, todas a informantes claves: 3 empleadores y 5 trabajadores por cuenta propia, los cuales se pertenecían a distintas actividades económicas.

La pauta aplicada en las entrevistas cognitivas contó con 24 preguntas, provenientes de 6 módulos distintos de un total de doce módulos de la encuesta. Para realizar las preguntas y sondeos, se hicieron 10 grupos, vinculando temas o conceptos medidos en el cuestionario piloto. Los grupos de preguntas por módulo y los temas o conceptos medidos fueron los siguientes:

**Tabla 16: Preguntas y Conceptos medidos en Entrevistas Cognitivas**

Preguntas	Conceptos	Módulo
<b>B1-B2</b>	Oportunidad, necesidad, complementar ingresos, expectativas de mayor ingreso	Módulo B
<b>D3</b>	Materias primas e insumos, arriendo de maquinaria, gastos del negocio	Módulo D
<b>D4-D5</b>	Actividades agrícolas, ganaderas, forestales y pesqueras, autoconsumo, productos derivados	Módulo D
<b>D6-D7-D8</b>	Actividades agrícolas, ganaderas, forestales y pesqueras, consumo propio, consumo del hogar, promedio mensual	Módulo D
<b>D9-D10-D11</b>	Equipos, vehículos, maquinarias o herramientas, ganancia neta	Módulo D
<b>D12-D13-D14-D15</b>	Persona socia, sueldo mensual, ganancia mensual	Módulo D
<b>E1-E2</b>	Contabilidad, registros personales, gastos del hogar, gastos del negocio	Módulo E
<b>I1-I2-I3</b>	Uso de internet, promocionar empresa, vender bienes y servicios, hacer trámites, pago de remuneraciones	Módulo I
<b>K3</b>	Proveedor, Cliente	Módulo K
<b>L3-L4-L5</b>	Cuidado, dependencia permanente, personas del hogar	Módulo L

Fuente: INE, VI Encuesta de Microemprendimiento, 2019

<sup>25</sup> Ver Anexo 2

## II. Resultados entrevistas cognitivas

A continuación, se presentan algunos de los resultados encontrados según los conceptos mencionados y los grupos de preguntas en cuestión, con la intención de facilitar el análisis y entregar un valor que distinga a esta técnica respecto a las otras presentadas a lo largo del piloto de la VI EME.

### **Módulo B – Historia Laboral**

El principal problema tuvo relación con las diferencias en la comprensión de los conceptos de necesidad y oportunidad entre tipos de informantes. En consecuencia, algunas categorías de respuesta que se clasificaron en el grupo de “necesidad” eran vistas por los informantes como una “oportunidad” y viceversa.

### **Módulo D – Ingreso, gastos y ganancias del negocio.**

Se evidenció dificultad en la comprensión de los conceptos de materias primas e insumos, la diferencia entre gastos en servicios profesionales o remuneraciones y la diferencia entre los ingresos y ganancias. Además, se observaron problemas para recordar y estimar ciertos gastos y la periodicidad de ellos.

### **Módulo L- Trabajo no remunerado**

En las preguntas sobre trabajo de cuidados, algunos informantes tuvieron problemas para comprender si la respuesta apunta al tiempo relacionado a la actividad por cuenta propia o al tiempo libre de los y las informantes que trabajan en forma remunerada cuidando a otros. además, se observaron problemas de estimación y respuesta para calcular el tiempo dedicado al cuidado de otras personas.

#### 5.3.2. Observaciones Técnicas

Una de las técnicas utilizada por más tiempo en la Encuesta de Microemprendimiento es la observación técnica. Esta metodología permite observar en terreno parte de los problemas presentes en el proceso pregunta-respuesta de los encuestados.

#### I. Metodología observación técnica

Este proceso se realiza a partir de un acompañamiento al equipo de encuestadores en terreno, donde se busca observar durante distintas jornadas de trabajo (e idealmente en distintos perfiles de informantes) cómo es el proceso de aplicación del cuestionario. La metodología implica que la persona que realiza la observación no puede interrumpir el trabajo de campo del equipo operativo<sup>26</sup> sino más bien deba tratar de observar cómo se desenvuelven las partes durante una jornada habitual de trabajo.

---

<sup>26</sup>Sólo en casos donde el observador técnico detecte errores graves en el levantamiento del cuestionario, se permite dar una voz de alerta al encuestador para rectificar el procedimiento en el momento oportuno.

Para sistematizar los hallazgos, se diseñó un cuadernillo de campo<sup>27</sup> que permitió transcribir de forma ordenada los principales hallazgos observados en la entrevista. El objetivo principal de la planilla fue llevar un registro de las principales situaciones críticas por tipo de hallazgo. Para ello se consideró los mismos tipos de hallazgos provenientes del informante junto con algunos hallazgos relacionados al cuestionario o encuestador.

En el contexto de la prueba piloto de la EME VI se realizaron 14 observaciones técnicas. La aplicación del cuadernillo de campo estuvo a cargo del equipo técnico de la encuesta y contó con el apoyo del equipo de operaciones. El levantamiento se realizó en zonas urbanas y rurales, tanto de las comunas de la RM como en la región de O'Higgins.

## II. Resultados observación técnica

Se registraron 110 hallazgos en total, dentro de los cuales el más recurrente fue el de comprensión seguidos por los problemas de estimación. El detalle de éstos se puede observar en la siguiente tabla.

**Tabla 17: número de hallazgos por cada tipo**

Tipo de hallazgo							
Comprensión	Memoria	Estimación	Respuesta	Secuencia	Parfraseo inadecuado	Otro	Total
42	16	22	18	1	2	9	110

Fuente: INE, VI Encuesta de Microemprendimiento, 2019

### 5.3.3. Mesas de conversación

En el período de levantamiento de la prueba piloto se realizaron algunas mesas de conversación con el equipo operativo. Esta técnica de evaluación consiste en una discusión grupal en torno a temáticas específicas, conversación que es guiada por un por un miembro del equipo técnico o de operaciones, y donde se pretende indagar acerca de la percepción y opinión de los asistentes frente a un fenómeno, siendo en este caso la aplicación del cuestionario en terreno.

Las mesas de conversación son similares a un grupo focal, buscando responder interrogantes de todos los participantes en el marco de la interacción del grupo, permitiendo recolectar información relevante y pertinente para el mejoramiento del cuestionario, el manual de trabajo de campo, las capacitaciones y los procedimientos de análisis.

#### I. Metodología Mesas de conversación

El objetivo general de las mesas de conversación fue evaluar el proceso de levantamiento piloto de la VI Encuesta de Microemprendimiento 2019 desde la perspectiva de los miembros del equipo operativo y de trabajo de campo (supervisores, analistas y encuestadores).

<sup>27</sup>Anexo 1: Cuadernillo de campo

La técnica aplicada permite recoger y analizar las dinámicas y los conceptos construidos de manera grupal a partir de la experiencia de aplicación de la encuesta en terreno. Esta técnica cualitativa se encuentra estructurada mediante una pauta que orienta el diálogo sin limitar el desarrollo de la conversación, dando flexibilidad a que surjan nuevos elementos para el proceso. La pauta para la mesa de conversación aborda distintas temáticas relacionadas en primer lugar a la recepción del informante, luego aborda algunas dificultades del encuestado presenciadas por el encuestador, clasificándolas en los tipos de hallazgos definidos en el apartado de entrevistas cognitivas (comprensión, memoria, estimación y respuesta), finalmente se detallaban algunos problemas del cuestionario

## II. Resultados Mesas de conversación

A continuación, se muestran los resultados obtenidos de las tres mesas de conversación llevadas a cabo. La pauta para guiar esta metodología será la utilizada para exponer los resultados obtenidos.

### **Recepción del informante en terreno**

El principal aspecto destacado en este ítem fue la larga duración del cuestionario, la que en promedio fue de 50 minutos aproximadamente, 20 minutos más que en su versión anterior. Esto dificultó el desarrollo normal de la aplicación del cuestionario puesto que los y las informantes manifiestan cansancio y distracción, generalmente después de 30 minutos de aplicación.

### **Claridad de las preguntas (comprensión)**

En relación con la comprensión de las preguntas del cuestionario, existieron diversos problemas los cuales en su mayoría se encontraban también en el módulo D. Ingreso, gastos y ganancias del negocio.

### **Memoria (recuperación de información)**

Uno de los módulos que presentó más dificultades en la recuperación de información fue el de “Gastos, ingresos y ganancias del negocio”, principalmente en la pregunta referida a los gastos del negocio (. Según el equipo, la mayor complicación se presentó debido a la extensión del período de referencia indagado (durante el año 2018).

### **Estimación de respuestas**

Según lo indicado por los encuestadores las preguntas con mayores problemas de estimación fueron las del módulo D “Gastos, ingresos y ganancias del negocio”, lo cual se relacionó con la baja comprensión de algunos conceptos.

### **Flujo y forma de registro**

En general, las y los encuestadores manifiestan problemas de llenado en los módulos “Gastos, ingresos y ganancias del negocio” y “Trabajo no remunerado”.

#### 5.3.4. Análisis Base de datos

La metodología del análisis de la base de datos permite identificar la prevalencia, incidencia, asociación y correlación entre ciertas variables, información que fue de gran utilidad para evaluar el aporte de las preguntas y la consistencia de la información utilizada.

Los análisis estadísticos desarrollados permitieron identificar posibles mejoras a las preguntas, las cuales se traducen en la eliminación o fusión de preguntas, y en la inclusión, reordenamiento o eliminación de alternativas. En esta sección se detalla la metodología y los análisis que sirvieron como antecedentes en las modificaciones aplicadas.

#### III. Metodología análisis Base de datos

El análisis de la estructura de los datos corresponde a la descripción de las tendencias centrales para las variables de carácter continuo (medias, medianas y desviación). Para ello se utilizó la construcción de tablas de resumen con esta información estadística, mientras que para el caso de variables categóricas se realizaron tablas de frecuencias relativas.

El análisis de los datos se centró en las variables principales de los módulos D, G, L, debido a que las preguntas presentes en dichos módulos contienen información fundamental para la utilidad de la encuesta.

La programación de tablas de frecuencias, el procesamiento descriptivo y programación se realizó en el programa estadístico STATA.

#### IV. Resultados análisis Base de datos

Módulo D – Gastos, ingresos y ganancias del negocio.

Al analizar los resultados de las preguntas de gastos ingresos y ganancias se pudieron apreciar diferencias en la participación e incidencia de distintos ítems sobre el monto total de los gastos o ingresos.

En cuanto a la pregunta de gastos, los que concentraron la mayor prevalencia fueron las de “materias primas e insumos”, “servicios básicos” y “gasolina y diésel”.

Por otra parte, el gasto del ítem “cuotas o intereses por préstamos del negocio”, se pudo apreciar que esta información se capturó mejor a través de la grilla que desagregaba el tipo de deuda que mediante la pregunta directa.

Módulo L – Trabajo no remunerado

Al observar los resultados del módulo de trabajo no remunerado, a nivel general se apreció una mayor prevalencia e incidencia en la participación y carga de trabajo no remunerado en las mujeres.

En términos descriptivos, el trabajo no remunerado resultó ser bastante prevalente en las personas microempendedoras, siendo mayor en mujeres que en hombres, y en muchos casos representa más tiempo a este tipo de trabajo que al trabajo en la ocupación. En este sentido, se justificó la permanencia del módulo de trabajo no remunerado dada las formas en como las personas realizan sus actividades y su relación con las actividades por cuenta propia que llevan a cabo.

#### **5.4. Método de capacitación para la recolección de datos oficial**

La VI versión de la EME tuvo dos procesos de capacitación, uno orientado hacia los encuestadores y supervisores, realizado tanto en regiones como en la RM, y otro proceso direccionado hacia los analistas del equipo de operaciones del INE quienes fueron los encargados de realizar una primera validación de la base de datos de la encuesta. A continuación, se comentan algunos de los aspectos centrales de ambos procesos.

##### **5.4.1. Capacitaciones encuestadores y supervisores**

Estas capacitaciones tuvieron como objetivo brindar a los participantes los conocimientos técnicos y operativos para el correcto levantamiento del estudio. Además, se les entregó información sobre la Ley N° 17.374, que da cuenta de la confidencialidad de la información del concepto de secreto estadístico que se aplica a cualquier producto del INE.

En términos de tiempo, la capacitación se realizó a lo largo de tres días y medio, donde la primera y la última media jornada estuvo a cargo del equipo operativo, mientras que las demás jornadas estuvieron a cargo del equipo técnico.

Desde el equipo operativo, los relatores abordaron todos los temas relacionados con el trabajo de terreno, protocolos de levantamiento y aspectos administrativos. Se presentó, además, en este primer día de capacitación, los antecedentes del proyecto, los conceptos de terreno, el rol del encuestador, la Hoja de Ruta y la Tarjeta de Registro del Hogar.

Desde el equipo técnico, los profesionales realizaron una revisión exhaustiva de todas las preguntas del cuestionario, explicando el objetivo de cada una de ellas y entregando las herramientas conceptuales que pretendían ayudar a los encuestadores a enfrentarse a situaciones complejas en terreno, implementando ejemplos de posibles situaciones a enfrentar y soluciones efectivas a estas mismas.

En lo que respecta a la selección y reclutamiento de los recursos humanos necesarios para cumplir con los objetivos operativos del proyecto, se realizaron contrataciones en todas las regiones del país, llegando a un total de 52 encuestadores y 19 supervisores a nivel nacional, todos ellos bajo la responsabilidad de un encargado regional, que en la mayoría de los casos correspondió a la jefatura operativa de las Direcciones Regionales del INE. Adicionalmente, se conformó un equipo de profesionales encargados de realizar las actividades de gabinete, relacionadas con la recepción de los cuestionarios, la digitación de encuestas, la revisión de los cuestionarios en papel, y el desarrollo informático del sistema de ingreso de datos.

Se hizo entrega de los siguientes materiales de apoyo a los encuestadores al inicio de la capacitación:

- Manual de trabajo de campo
- Tarjetero en tamaño carta horizontal
- Tarjeta de Registro del Hogar
- Cuadernillo de Ejercicios

Además, durante la capacitación se desarrollaron ejercicios de juegos de roles con el fin de recrear situaciones reales que pueden suceder en terreno, dando mayor dinamismo a las capacitaciones entregadas. Al final de cada jornada de capacitación se realizó una prueba para ir monitoreando el aprendizaje de los encuestadores y supervisores, y reforzar áreas débiles en las siguientes jornadas.

En el caso de los supervisores, si bien asistieron a la capacitación en conjunto con los encuestadores, se les realizó una inducción específica que tuvo por objetivo darles a conocer los principales criterios de supervisión establecidos para el proyecto, instruyéndolos en el uso del sistema de gestión y en sus labores, tanto de terreno como de gabinete.

#### 5.4.2. Capacitación analistas

La capacitación entregada a los analistas de operaciones, quienes son los responsables de identificar los errores e inconsistencias en los datos registrados y de aplicar los procedimientos de depuración establecidos.

Este instructivo se realizó en Santiago en las oficinas del INE, Las principales temáticas abordadas en esta jornada dicen relación con los conceptos principales que debían manejar para comprender los procesos vinculados al control de calidad de los datos.

### **5.5. Supervisión**

#### 5.5.1. Supervisión directa

La supervisión directa es aquel proceso de evaluación *in situ* del trabajo del encuestador, es decir, que se realiza mientras este aplica el cuestionario al informante. Este proceso permitió al supervisor y al equipo técnico a cargo de la EME identificar rápidamente problemas que podrían estar experimentando algunos encuestadores, facilitando la oportuna realización de ajustes en el trabajo, los que pueden llegar incluso la reinstrucción del encuestador.

#### 5.5.2. Supervisión indirecta

La supervisión indirecta consistió en la aplicación de una pauta, que buscaba corroborar el protocolo de entrevista ejecutado por la persona encuestadora. Adicionalmente, se incluyeron preguntas enfocadas en corroborar la situación ocupacional de los informantes seleccionados.

### Figura 3: Pauta de evaluación indirecta.

1. A1. Usted dijo haber trabajado como Empleador o Trabajador por Cuenta Propia en \_\_\_\_\_ de 2019. ¿Continúa con esa actividad?

Sí, la misma actividad económica	No, sin embargo tengo otra actividad económica como Empleador o Cuenta Propia	No, cambié de situación laboral
1	2	3
<b>PASE A PREGUNTA 4</b>		

2. El/la encuestador/a, ¿le preguntó por qué motivo terminó su trabajo como cuenta propia?

Sí	No
1	2

3. El/la encuestador/a, ¿le preguntó si volvería a trabajar como cuenta o propia, o a qué se dedica actualmente?

Sí	No
1	2
<b>FIN SUPERVISIÓN</b>	

4. Al momento de la encuesta, ¿tenía empleados en su negocio o para la actividad que realiza?

Sí	No
1	2

5. El/la encuestador/a, ¿le preguntó cuál era su ocupación y la actividad principal de su negocio?

Sí	No
1	2

6. El/la encuestador/a, ¿le preguntó si ha solicitado un crédito para fines del negocio, durante los últimos dos años?

Sí	No
1	2

7. El/la encuestador/a, ¿preguntó si realiza distintas actividades no remuneradas o de cuidado de personas?

Sí	No
1	2

8. ¿Cuánto diría usted que duró la entrevista?

Menos de 10 minutos	Entre 10 y 20 minutos	Más de 20 minutos	NS / No Recuerda
1	2	3	88

Fuente: INE, VI Encuesta de Microemprendimiento, 2019



## 6. METODOLOGÍA DE PROCESAMIENTO DE DATOS

En el presente apartado se detallan las metodologías del procesamiento de datos y de las bases

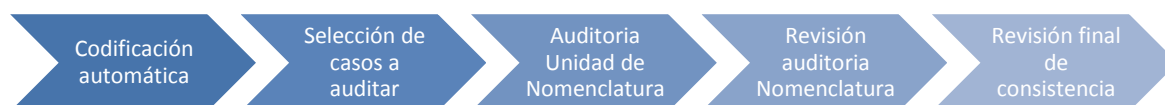
### 6.1. Codificación y clasificación

El trabajo de codificación y clasificación para esta versión de la Encuesta de Microemprendimiento contó con dos aristas, la rama de actividad económica y la ocupación u oficio del informante. A ello se sumó el trabajo de codificación de las preguntas con opciones “Otro, especifique”.

#### 6.1.1. Codificación CAENES y CIUO

La codificación tanto de CAENES<sup>28</sup> como de CIUO<sup>29</sup> se hizo tras la realización de cada una de las tres submuestras en las que se dividió la VI EME. La secuencia de trabajo contó de cinco fases sucesivas que se iniciaban con el proceso de codificación automática (CA) realizado por la unidad de procesamiento de la Encuestas Nacional de Empleo (ENE), tras lo cual el equipo técnico de la EME seleccionó una parte de los registros para realizar una auditoría que verificara la calidad de la CA. La unidad de auditoría revisó las glosas enviadas, clasificándolas de manera tal que pudiera corroborar que el software que ejecutó la CA realizó una asignación correcta. En el caso que ambos códigos, el de la auditoría y el de la CA fuesen distintos, las glosas eran revisadas nuevamente por el equipo de procesamiento de la ENE para determinar la glosa correcta. Finalmente, se realizó un análisis de consistencia de las glosas con otras preguntas del cuestionario.

#### Diagrama 6: Proceso de codificación



Fuente: INE, VI Encuesta de Microemprendimiento, 2019

#### I. Proceso de codificación automática

El proceso de codificación automática que se usó en la VI EME fue el Support Vector Machine (SVM). Este modelo es el mismo utilizado por la Encuesta de Empleo <sup>30</sup>.

A partir de las variables de la actividad principal del negocio y del oficio u ocupación se codificaron automáticamente la Rama de Actividad Económica a nivel de división, según el Clasificador de Actividades Económicas Nacional para Encuestas Sociodemográficas

<sup>28</sup> Adaptación nacional para encuestas sociodemográficas de la cuarta versión del Clasificador Internacional Industrial Uniforme (CIIU rev. 4) de la división de estadísticas de las Naciones Unidas

<sup>29</sup> Adaptación nacional del Clasificador Internacional Uniforme de Ocupación (CIUO) de la Organización Internacional del Trabajo (OIT) en su última 2008.

<sup>30</sup> [https://www.ine.cl/docs/default-source/ocupacion-y-desocupacion/metodologia/espanol/documento-sistema-de-clasificaci%C3%B3n-y-codificaci%C3%B3n-autom%C3%A1tica-\(mayo-2019\).pdf?sfvrsn=ceea6423\\_3](https://www.ine.cl/docs/default-source/ocupacion-y-desocupacion/metodologia/espanol/documento-sistema-de-clasificaci%C3%B3n-y-codificaci%C3%B3n-autom%C3%A1tica-(mayo-2019).pdf?sfvrsn=ceea6423_3)

(CAENES<sup>31</sup>) y el Oficio (Grupo Ocupacional) a nivel de subgrupo principal según el Clasificador Chileno de Ocupaciones (CIUO 08.CL<sup>32</sup> en adelante CIUO).

## II. Selección de casos a auditar

Una vez codificadas las glosas a través del proceso de codificación automática se seleccionaron los casos a auditar. Esta selección consideró a todos los casos que indicaron que si bien seguían trabajando como independientes cambiaron su actividad económica y a los casos que mantuvieron su actividad económica pero la actividad económica declarada, difería de la declarada en la ENE.

Al comparar las glosas entre la codificación automática de la EME y la de la ENE se pudo apreciar que la mayor parte de los casos presentan la misma clasificación tanto en CIUO como en CAENES.

**Tabla 18: Comparación de codificación ENE- EME**

Glosa codificar	a	Submuestra 1 y 2		Submuestra 3		Estado
		Frecuencia	Porcentaje	Frecuencia	Porcentaje	
<b>Sin diferencias</b>		3.500	67,3%	1.750	66,6%	Casos sin problemas
<b>Diferencias en CAENES</b>		532	10,2%	227	8,6%	Casos para auditar
<b>Diferencias en CIUO</b>		833	16,0%	389	14,8%	
<b>Diferencias en ambas</b>		335	6,4%	262	10,0%	
<b>Total</b>		5.200	100,0%	2.628	100,0%	

Fuente: INE, VI Encuesta de Microemprendimiento, 2019

### 6.1.2. Codificación y reclasificación específica

La codificación de las respuestas literales entregadas por los informantes en aquellas preguntas que contaban con la opción “Otro” en el cuestionario fue la otra gran tarea de codificación y clasificación que se realizó para esta versión de la EME.

El cuestionario incluyó 22 preguntas que tenían entre sus alternativas la posibilidad de responder “Otro”, ante lo cual se le pidió al informante que especifique su respuesta. Se decidió categorizar las respuestas literales entregadas por los encuestados generando categorías generales que permitieran reducir el volumen de información. Este proceso

<sup>31</sup> [https://www.ine.cl/docs/default-source/ocupacion-y-desocupacion/publicaciones-y-anuarios/publicaciones/clasificador-de-actividades-econ%C3%B3micas-nacional-para-encuestas-sociodemogr%C3%A1ficas-\(caenes\)c9163cea683e61618024ff030098b759.pdf?sfvrsn=942ae665\\_3](https://www.ine.cl/docs/default-source/ocupacion-y-desocupacion/publicaciones-y-anuarios/publicaciones/clasificador-de-actividades-econ%C3%B3micas-nacional-para-encuestas-sociodemogr%C3%A1ficas-(caenes)c9163cea683e61618024ff030098b759.pdf?sfvrsn=942ae665_3)

<sup>32</sup> [http://webanterior.ine.cl/docs/default-source/publicaciones/2018/ciuo-08.cl-clasificador-chileno-de-ocupaciones.pdf?sfvrsn=50bd58d2\\_4](http://webanterior.ine.cl/docs/default-source/publicaciones/2018/ciuo-08.cl-clasificador-chileno-de-ocupaciones.pdf?sfvrsn=50bd58d2_4).

incluyó, además, reasignar algunas respuestas que fueron catalogadas como “Otro” pero que deberían haber sido consideradas dentro de las alternativas principales de la pregunta de origen.

## **6.2. Análisis de consistencia y edición**

El trabajo de consistencia y edición se dividió en dos grandes ramas: la aplicación de una serie de reglas de validaciones y el análisis y corrección de valores atípicos. En los siguientes párrafos se detalla las tareas realizadas en ambos ámbitos.

### **6.2.1. Diseño e implementación de validaciones**

Para controlar la calidad de la información levantada, el equipo técnico diseñó e implementó una serie de reglas de validación de la información con el fin de que las unidades encuestadas tuvieran completitud en la información pesquisada y pudieran analizar la consistencia de los datos recolectados.

#### **I. Diseño de reglas de validación**

El equipo técnico diseñó una planilla de validación que entregó a la contraparte operativa para que, una vez digitados los formularios, pudieran identificar dos tipos de inconsistencias: primer y segundo nivel. Las validaciones de primer nivel permitieron identificar problemas de flujo y rango para luego solucionar estos casos, es decir, cuando un informante no respondió una pregunta, presentó una respuesta sin tener que hacerlo o bien la respuesta entregada se encontraba fuera del rango de alternativas posible.

Estas validaciones tenían dos procedimientos de corrección: la devolución a terreno y la edición de datos en base a alguna metodología de depuración previamente definida.

#### **II. Implementación de reglas de validación diseñadas**

Tras la realización de cada submuestra, la Subdirección de Operaciones envió al equipo técnico dos bases de datos denominadas base full y base de empleo. Sobre estas bases, el equipo técnico ejecutó la planilla de validación para identificar si todavía existían problemas de primer y segundo nivel que no hubiesen sido identificados por la contraparte operativa.

En relación con los problemas de segundo nivel, se diseñaron 85 validaciones que tenían como objetivo corregir problemas de coherencia global de la encuesta. Es importante indicar que en la mayor parte de los validadores de segundo nivel el proceso de corrección consistía en recontactar a los informantes, lo que se realizó en gran parte de los casos. Para las inconsistencias que no pudieron ser resueltas a través de este mecanismo o que podían ser abordadas de otra manera, se diseñaron reglas de depuración.

#### **III. Depuración de la base de datos**

Además de las inconsistencias programadas antes del trabajo en terreno, en los procesos de recolección y procesamiento, se detectaron algunos imprevistos adicionales para los cuales se diseñaron reglas de validación. Como se comentó anteriormente el primer paso

fue contactar a los encuestados para validar la información otorgada, no obstante, en algunos casos no fue posible esto por problema de contacto de los informantes o bien limitaciones operativas. Para enfrentar este tipo de situaciones, se realizaron algunas depuraciones cuyo objetivo era mejorar la consistencia y completitud de los datos.

a) Depuración de casos Fuera de Marco

Este problema se presentó cuando se identificaron informantes que no debieron haber sido seleccionados para participar en el estudio debido a que no eran trabajadores independientes, no eran residentes habituales de las viviendas seleccionadas o eran trabajadores independientes dueños de una unidad de mayor impacto (más de 10 trabajadores). Para ello se recurrió a la información disponible en los módulos del cuestionario y a las observaciones de cada caso.

b) Depuración automática de inconsistencias

Con el objetivo de que los datos de todas las variables fueran válidos, se realizó un análisis en base a las reglas de validación establecidas que incluía la revisión de observaciones y la depuración automática de las observaciones que tenían una solución evidente. Ejemplo de estos casos es el siguiente:

- **Boleta de Honorarios e IVA:** Esta inconsistencia se presentó en informantes que declaraban IVA (e6\_1=1) y haber iniciado actividades en el SII como trabajador/a independiente con boleta de honorarios (e3=1) lo cual es una inconsistencia ya que los servicios no estaban grabados con IVA. El procedimiento de corrección adoptado fue conservar la respuesta en e3 sobre el inicio de actividades y modificar la respuesta sobre la declaración de IVA, cambiando a “no aplica” (e6\_1=85) y que “sí” realizó trámites de declaración de impuesto a la renta (e6\_2=1).

c) Identificación de valores atípicos

La identificación de los valores atípicos se aplicó a todas las variables continuas mediante el modelo estadístico Bacon (blocked adaptive computationally efficient outlier nominators) y correspondiente análisis de casos. Se excluyó de este análisis a las variables de trabajo no remunerado, las que tuvieron un procedimiento de depuración distinto creando variables editadas que contenían un tratamiento a los valores atípicos.

En la sección siguiente se ahonda en la identificación, análisis y tratamiento de estos casos.

6.2.2. Revisión valores atípicos

A partir de una evaluación de las metodologías de detección de outliers<sup>33</sup> a través de la variable de ganancia anual (D7\_2) el equipo técnico de la VI EME decidió trabajar con el modelo estadístico *bacon* (blocked adaptive computationally efficient outlier nominators)

---

<sup>33</sup> Para mayores detalles revisar “Documento de tratamiento de valores atípicos en VI EME”.

Además, se decidió que los perfiles de informantes se determinarían a través de la combinatoria de 4 variables:

- Sexo: variable nominal que asume los valores de hombre (1) o mujer (2)
- Número de trabajadores/as asalariados: variable de razón que indica la cantidad total de asalariados/as en la unidad económica, siendo 0 el valor mínimo esperado.
- Registro en el SII: variable nominal con 6 categorías de tipo de registro en el Servicio de Impuestos internos
- CAENES: variable nominal que identifica el tipo de actividad a la que se dedica la unidad económica en la que trabaja la o el informante.

En primera instancia, para determinar los perfiles se consideraron también, las variables edad (cuantitativa), nivel educacional (ordinal) y antigüedad del negocio (de menos de un año a 10 años o más). Sin embargo, fueron descartadas porque limitaban los casos iniciales que cumplían los requisitos de tener respuesta válida en todas las variables. Además, se decidió dejar fuera estas variables pues, en algunos casos ingresaba información redundante al modelo observándose colinealidad, lo que generaba que el modelo no presentara resultados gestionables.

#### I. Casos outliers por preguntas cuantitativas

Con el método de *bacon* y considerando las cuatro variables independientes detalladas en los puntos precedentes, se realizó el conteo general de casos fuera de rango para todas las variables cuantitativas a analizar en cada una de las submuestras. Para las variables que podrían ser reportadas con distinta periodicidad (d3, d5, d6), se anualizaron todos los valores y desde esa nueva variable se realizó la búsqueda de los casos fuera de rango. De esta manera, se corrigió el problema de las distintas periodicidades. A modo de ejemplo, a continuación, se muestra el detalle de los outliers encontrados en la submuestra 1.

**Tabla 19: casos totales y casos outliers**

Variable	N	outliers	%	Mínimo valor outlier
B9 sueldo líquido	777	4	0,5%	4.000.000
D1 pago arriendo	350	3	0,9%	4.500.000
D2 estimación arriendo	593	3	0,5%	6.000.000
D3c_1 gasto en materias primas e insumos	1750	50	2,9%	100.000
D5d_1 ingreso por cultivo	187	10	5,3%	15.000.000
D6c ingreso promedio	1943	33	1,7%	4.500.000
D7_1 ganancias el mes anterior*	2211	39	1,8%	1.200.000
D7_2 ganancias los últimos 12 meses*	1923	32	1,7%	10.000.000
D12_monto otros ingresos	871	9	1,0%	290.000
G9_1 cuota mensual crédito bancario	279	6	2,2%	5.000.000
H3_1 valor de computadores, celulares, etc	1461	14	1,0%	4.000.000
B9 sueldo líquido	777	4	0,5%	4.000.000

Nota: \* En d7\_1 también se identificaron algunos outliers en la cola inferior de la distribución, los cuales tomaron los valores de -5.000.000 y 0. En d7\_2, ocurrió el mismo fenómeno, tomando un valor de -60.000.000. Fuente: INE, VI Encuesta de Microemprendimiento, 2019

Es importante indicar que, los casos identificados como outliers presentaron un rango distinto en cada uno de los perfiles a analizar (que se construyeron a través de la combinatoria de las variables), lo que justifica la utilización del método *bacon* por sobre el establecimiento de un límite fijo para todos los informantes.

El cuadro precedente evidencia los diversos límites de casos fuera de rango que se obtuvieron para cada una de las variables cuantitativas analizadas, por lo que se sostiene la premisa de que el modelo es una mejor aproximación al total de casos outliers que la simple fijación de un límite uniforme para todas las variables.

## II. Comparación de los casos dentro y fuera de rango

Es importante recordar que el modelo *bacon* genera perfiles de sujetos a partir de la combinatoria de las variables del modelo, y para cada una de estas combinatorias o perfiles identifica valores fuera de rango. En este sentido, el valor mínimo de los outliers presentado

en la tabla 9 no funciona como un punto de corte, sino que es el menor valor obtenido considerando todos los perfiles generados a partir de las variables independientes, es decir que el valor mínimo proviene de la totalidad de los perfiles analizados, y no corresponde a un perfil en particular. Por ello, es posible que existan valores que estén por sobre dicho límite pero que no necesariamente sean considerados como casos fuera de rango, debido al perfil del informante que entrega ese dato.

A pesar de ello, en la mayoría de los casos se observan importantes diferencias entre los valores en rango y los valores outliers, lo que quedó de manifiesto cuando se compararon las medias de ambos escenarios, como lo presenta la siguiente tabla.

**Tabla 20: Medias para casos en rango y casos outliers**

Variable	Media casos en rango sin outliers	Media solo casos outliers	Min. outlier
B9 sueldo líquido	506.244	4.625.000	4.000.000
D1 pago arriendo	366.889	7.166.667	4.500.000
D2 estimación arriendo	409.118	33.700.000	6.000.000
D3c_1 gasto en materias primas e insumos	259.345	11.600.000	100.000
D5d_1 ingreso por cultivo	871.601	51.400.000	15.000.000
D6c ingreso promedio	461.191	19.300.000	4.500.000
D7_1 ganancias el mes anterior*	413.598	6.904.017	1.200.000
D7_2 ganancias los últimos 12 meses**	4.673.461	107.000.000	10.000.000
D12_monto otros ingresos	190.864	2.115.556	290.000
G9_1 cuota mensual crédito bancario	398.440	18.300.000	5.000.000
H3_1 valor de computadores, celulares, etc	290.801	7.714.286	4.000.000
H4_monto inversión en herramientas	2.164.548	44.400.000	10.000.000

Fuente: INE, VI Encuesta de Microemprendimiento, 2019

### III. Tratamiento valores outliers

El procedimiento que se adoptó ante los casos fuera de rango para asegurar la veracidad del dato en el caso que corresponda, o su reemplazo cuando sea pertinente constó de cuatro pasos:

1. Para cada caso identificado como outlier a través del método bacon, se revisaron todas las observaciones. Si el caso fuera de rango contaba con una observación que justificara el valor dado, entonces el dato se mantuvo en la base. Además, se solicitó la encuesta en papel para verificar si el dato fue correctamente digitado. En caso de que no se justificara por algunos de los motivos anteriores, se avanzó al segundo paso.
2. Se revisó si el caso contaba con algún otro valor atípico. Se asumió que existía coherencia si es que el caso presentaba dos o más datos outliers en el mismo módulo del cuestionario. De ser así, se mantenían los valores registrados en la base de datos. En caso contrario, se continuaba con el tercer paso.
3. Se evaluó si contaba con un valor atípico en las variables de ingreso anual, gasto y ganancia anuales al percentil 30 y de ser así, se validó el dato. Este procedimiento se fundamentaba en el hecho de que existen muy pocos casos que tenga más de un valor outlier gracias a que el nivel de exigencia de detección de valores atípicos es muy alto, por lo que fue necesario flexibilizar el criterio para solucionar algunos casos adicionales.
4. Finalmente, si el caso no satisfacía las condiciones impuestas por los tres pasos anteriores, se consideró que el dato no podía ser confirmado, por lo que se reemplazó la respuesta entregada por el informante por un valor 96, es decir, "sin dato". Esta decisión se basó en la imposibilidad de corroborar la veracidad de que el dato fuera de rango, y en el hecho que es coherente con el tratamiento de casos fuera de rango que realizan las encuestas ESI y ENUT de la Institución.

#### **6.3. Imputación y creación de variables**

Las imputaciones en este estudio se centraron en la triada que forman los gastos, ingresos y ganancias de los micronegocios, los que se recogieron en preguntas que experimentaron cambios respecto a la V EME. Este apartado también revisa el proceso de creación de los indicadores de trabajo no remunerado, una de las novedades de esta VI versión de la EME.

##### 6.3.1. Depuración y estimación de gastos, ingresos y ganancias

La VI EME indaga por los ingresos, los gastos y las ganancias de forma independiente, lo que permite contar con la ganancia declarada por el informante además de la ganancia estimada a través de las otras variables. En este contexto, se observaron cuatro problemas de ausencia de respuesta los cuales se describen a través de la siguiente tabla donde las casillas blancas corresponden a casos con respuesta válida y las negras a omisión de información.



**Tabla 21: Ausencia de respuesta en preguntas de gastos, ingresos y ganancias**

Ingreso	Gastos	Ganancias estimadas	Ganancias declaradas	Situación
				Caso 1
				Caso 2
				Caso 3
				Caso 4

El caso 1 describe a un informante con respuestas en todas las preguntas en estudio, donde además de tener una declaración de ganancias (responde pregunta d7\_2), fue posible calcularla. El caso 2 muestra informantes con respuestas en las preguntas de ingresos y gastos mas no en las de ganancias, sin embargo, fue posible estimar la ganancia a partir de la diferencia entre los ingresos y gastos. En el caso 3 se representan aquellos casos donde no es posible estimar una ganancia dado que no se contó con la información necesaria para hacerlo, pero que cuentan con declaración de ganancia por parte de los informantes. Por último, el caso 4 corresponde a informantes que no declararon la ganancia y ni tampoco fue posible estimarla.

A la ausencia de información en las preguntas de ingresos, gastos y ganancias se suma la existencia de inconsistencias en las respuestas detalladas en términos de la relación entre estas variables. Para poder solucionar estos problemas se diseñó una metodología de depuración e imputación de gastos, ingresos y ganancias<sup>34</sup> la cual se describe a continuación.

#### I. Caso 1: Corrección de inconsistencias en gastos, ingresos y ganancias

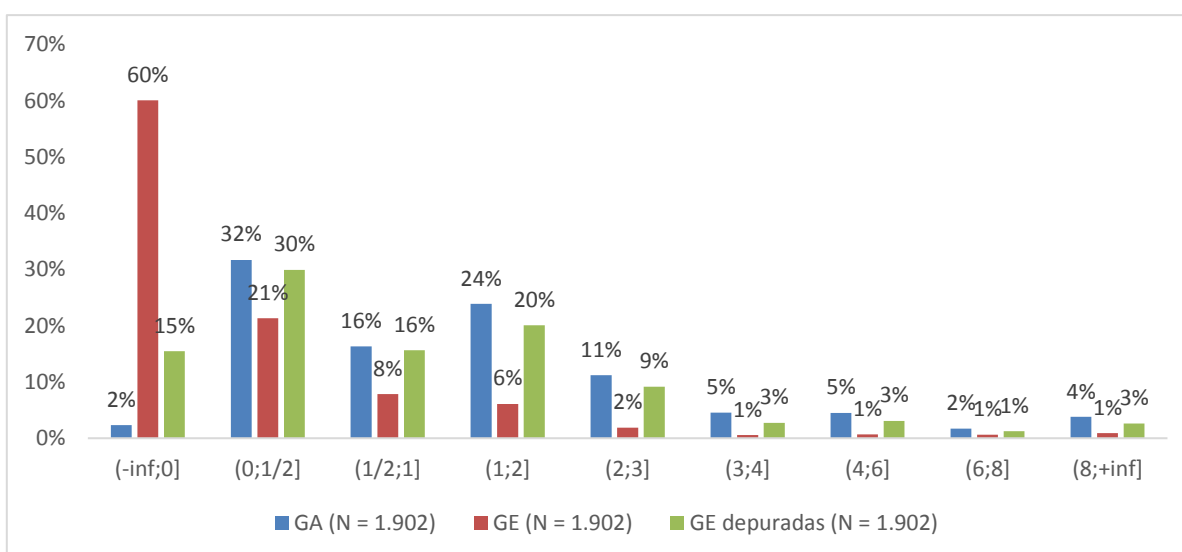
La depuración de datos fue una fase propia de la producción estadística que busca disminuir las inconsistencias lógicas. Este proceso fue realizado previo al proceso de imputación con el fin de mejorar la consistencia de la información de la ganancia del negocio para que la información fuese coherente y así poder contar con donantes de información para el siguiente proceso de imputación.

---

a) Depuración casos Ganancia estimada menor a ganancia declarada

Las visitas técnicas realizadas durante la aplicación del cuestionario revelaron que era frecuente que los informantes confundieran el concepto de ingresos del negocio con el de ganancias del negocio, declarando en la pregunta respecto a los primeros, la información de ganancias<sup>35</sup>. Por lo tanto, la depuración consistió en asignar a la variable de ganancia estimada el ingreso, y no calcular esta información a través de la resta de los ingresos y los gastos, aunque esto sólo se realizó en aquellos casos en que la nueva ganancia estimada no fuese mayor a la ganancia declarada por el informante.

**Gráfico 2: Distribución Ganancias estimadas antes y después de depurar (monto en sueldos mínimos)<sup>36</sup>**



Fuente: Elaboración propia. VI EME.

A nivel general, este proceso fue aplicado 1.902 informante, generando una distribución de frecuencia de ganancia depurada que se asemeja mucho a la distribución de la ganancia declarada por el informante como se aprecia en el gráfico anterior, disminuyendo el porcentaje de informantes que presentaban inconsistencias o ganancias estimadas negativas.

b) Depuración casos Ganancia estimada mayor a ganancia declarada

<sup>35</sup> A través de las visitas desarrolladas a terreno, se pudo observar que los conceptos de ingresos y ganancias tenían problemas de comprensión. Esta situación se reiteró al realizar la recuperación telefónica de información, apreciándose que al momento de contestar los ingresos algunos informantes descontaban los gastos directamente para realizar la estimación de la ganancia. Cabe señalar que el instructivo de los encuestadores incluía instrucciones para corregir esta confusión en terreno.

<sup>36</sup> Sueldo Mínimo desde el 1 de septiembre del 2018 al 28 de febrero del 2019 correspondió a \$288.000

Un segundo problema se produjo en aquellos casos en que la ganancia estimada, es decir aquella calculada en base a la información de ingresos y gastos entregada por los encuestados, era mayor que la ganancia declarada. Para las situaciones más críticas se decidió contactar a los informantes con el fin de analizar si la información entregada era correcta o había algún problema en el monto o la periodicidad reportada.

#### II. Caso 2: Estimación de ganancias en casos que no las declaran

Existió un 15,7% (1.154 casos) que no declaró ganancias, procediéndose a estimar las ganancias con información directa del encuestado por medio de las respuestas de ingresos y gastos del negocio. Para calcular la ganancia estimada depurada, al igual que en el caso 1, se evaluó si el informante confundió los conceptos de ingresos y ganancias. Para determinar si esto ocurrió, se analizó si la ganancia estimada es menor a cero o no. Además, se imputó una variable de ganancia donada<sup>37</sup> instrumental, la cual sólo fue utilizada para comparar la información entregada por el informante, que se comparó con la ganancia estimada y el ingreso declarado por el informante.

#### III. Caso 3: Sólo ganancia declarada

El caso 3 se trabajó de manera diferente a los demás, pues si bien se trata de informantes para quienes no es posible estimar una ganancia dado que no existía información necesaria para hacerlo (gastos e ingresos), sí existía una declaración de ganancia por parte de los informantes. En estos casos, esta ganancia declarada se mantuvo como la ganancia final del negocio.

#### IV. Caso 4: Imputación de ganancia anual

La última fase correspondió a la imputación de datos faltantes. Esta consiste en la aplicación de una metodología de donantes, la cual es utilizada en otras encuestas como la ESI<sup>38</sup>, CASEN, EPF, entre otras. En estos casos, se imputó la variable de ganancia final<sup>39</sup> por medio de un informante con las mismas características de aquel que no tenía información.

Los perfiles de informantes fueron elaborados en base a la experiencia de la ESI, la cual fue adaptada a la realidad de la encuesta. En este sentido se tomaron siete variables de la ESI (Actividad económica, sexo, edad, educación, horas trabajadas, situación en la ocupación y tamaño del negocio (según número de trabajadores) las cuales fueron recodificadas considerando las alternativas presentes en la encuesta. Además, se agregaron dos variables que indagan en la formalidad del negocio, las cuales a su vez están altamente relacionadas con las ganancias: tipo de inscripción ante el SII y registro contable.

A partir de estas variables se construyeron 13 niveles o perfiles de informantes, los cuales resultan de combinaciones de estas 9 variables. El primer nivel se construye a través de

---

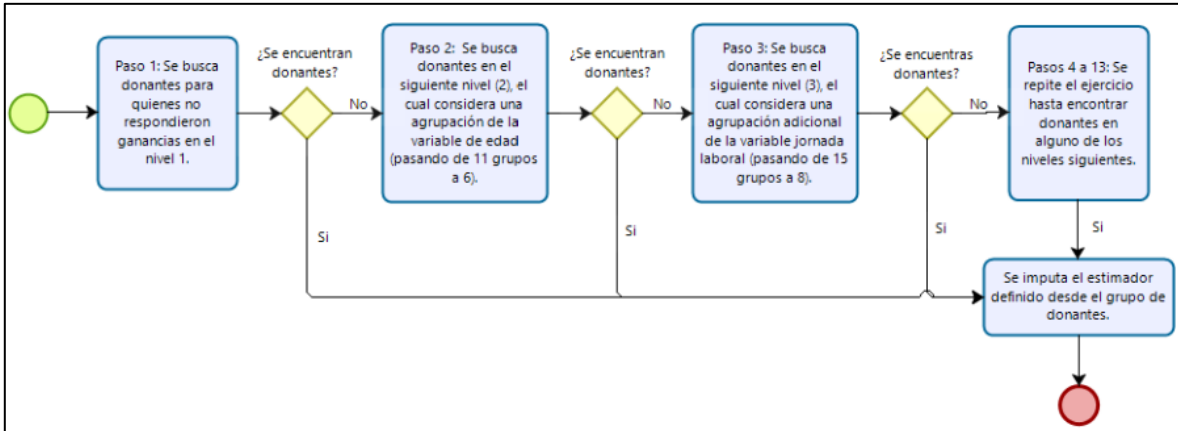
Esta variable se elabora en base a la metodología de donación de datos por medio de donantes o vecinos cercanos, es decir, aquellos informantes que sí responden la ganancia declarada le donan sus respuestas a informantes que no responden. Como estos informantes no responden la información de la ganancia declarada, esta donación se utilizará para evaluar si el informante confundió el concepto de ingresos con el de ganancias.

<sup>38</sup>[https://www.ine.cl/docs/default-source/ingresos-y-gastos/esi/ingreso-de-hogares-y-personas/antecedentes-metodol%C3%B3gicos/documento\\_metodologico\\_esi.pdf](https://www.ine.cl/docs/default-source/ingresos-y-gastos/esi/ingreso-de-hogares-y-personas/antecedentes-metodol%C3%B3gicos/documento_metodologico_esi.pdf)

<sup>39</sup> El proceso de construcción de esta variable se explica en detalle en el siguiente apartado

combinaciones desagregadas de estas 9 variables: las cuales presentan 18 categorías de actividad económica, 2 de sexo, 11 distintos grupos de edad, 6 niveles de educación, 16 jornadas de trabajo, 2 categorías ocupacionales, 5 de tamaño de la unidad económica, 7 tipos de Inscripción ante SII, 6 tipos de contabilidad.

**Figura 4: Proceso de búsqueda de donantes**

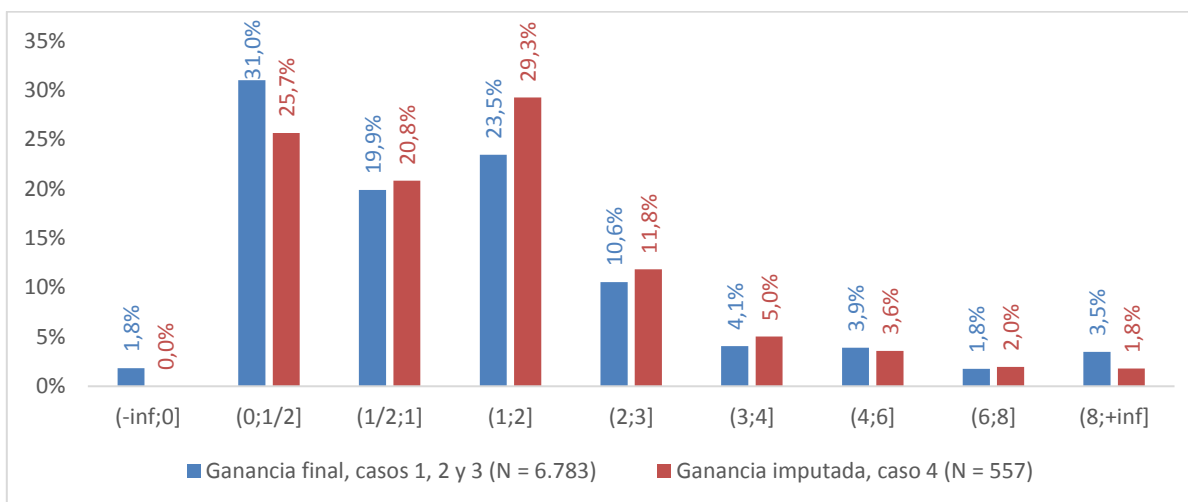


Fuente: Elaboración propia. VI EME.

El proceso de identificación de donantes comienza buscando en el primer nivel definido (ver Anexo 1) donantes que den su información de ganancias, y además que se hallaran en un área geográfica cercana. Si en esta área geográfica no se encontró a un donante de información, se pasó a un área geográfica más agregada. Las áreas geográficas siguieron el siguiente orden: Manzana, Estrato, Provincia, Región y País. Si no se encontraban donantes con el mismo perfil del receptor se avanzaba al siguiente nivel (Nivel 2), el cual fusiona ciertos grupos de edad (pasando de 11 a 6) y la presencia de contabilidad (pasando de 6 a 4). De no encontrar donantes en el nivel 2 en ninguna de las áreas geográficas se pasó al nivel 3, el cual considera una fusión en las variables de jornada de trabajo y de la inscripción en el SII, y así sucesivamente. El proceso terminó una vez encontrado donantes para todos los receptores de información.

Al aplicar el proceso de imputación se apreció una distribución bastante similar entre la ganancia final resultante de los casos 1, 2 y 3 y de la ganancia imputada en el caso 4. El detalle se puede observar en el siguiente gráfico:

**Gráfico 3: Distribución ganancia final (casos 1,2 y 3) y ganancia imputada (caso 4)<sup>40</sup>**



Fuente: Elaboración propia. VI EME.

Para esta versión de la encuesta se decidió aumentar el número mínimo de donantes de 1 a 3, debido a que este número minimiza el error cuadrático medio, mejorando la predicción de los datos faltantes. Esto se comprobó a través de un ejercicio explicado en el Anexo 1.

### 6.3.2. Construcción de variable ganancia final

Una vez depurada la variable de ganancia estimada se construyó la variable de ganancia final. Ésta toma los valores de la ganancia estimada depurada o la ganancia declarada dependiendo del caso. En líneas generales en los casos en que se contó con respuesta de gastos, ingresos y ganancias, la variable de ganancia final tomó el mayor valor entre ambas ganancias<sup>41</sup>.

En los casos que no se cuenta con información de ganancia declarada, se decidió que la ganancia final fuese igual a la ganancia estimada depurada, la cual tomó el valor de la ganancia estimada (Ingresos menos gastos) si los informantes no confunden los conceptos de ingreso y ganancia (Caso 2) y el ingreso en los casos que si confundan estos conceptos (Casos 2.).

En los casos que los informantes no contaban con respuestas de ingresos y gastos (lo que impide la construcción de la ganancia estimada) y sí respondan la ganancia declarada, la variable de ganancia final tomó el valor de la ganancia declarada (Caso 3). Finalmente, en el caso 4, se considerará el valor de la ganancia donada por medio de la metodología de medianas condicionadas descritas en el apartado anterior.

<sup>40</sup> Valores del eje X se encuentran expresados en sueldos mínimos

<sup>41</sup> Esta decisión se justifica debido a la subdeclaración de ingresos y/o ganancias presentes en las encuestas de ingresos. Este método también es realizado por el INEC de Costa Rica el cual considera la ganancia mayor cuando disponen de dos fuentes de información de ganancias del informante.

### 6.3.3. Visualización base de datos

Para poner a disposición el proceso de construcción de la “ganancia\_final” y facilitar el análisis a los usuarios, además de la información presente en el cuestionario, la Base Full contiene variables que permiten comprender el proceso de construcción de esta variable. Para ello, se incorpora las variables de gasto\_anual e ingreso\_anual las cuales son un insumo para generar las variables de ganancias estimadas (ganancia\_estimada y ganancia\_estimada\_depurada). También, se incorpora la variable de ganancia donada.

### 6.3.4. Generación variables trabajo no remunerado

Uno de los nuevos elementos presentes en la VI versión de la Encuesta de Microemprendimiento fue la incorporación de un módulo que recoge información respecto al trabajo no remunerado. Este módulo recoge información respecto a siete actividades o tareas relacionadas con el trabajo doméstico y el cuidado. A cada informante se le preguntó por la participación en cada una de estas actividades, cuántas veces a la semana las realizó y cuántas horas ocupó en dicha tarea la última vez que la realizó.

El proceso de creación de las variables se inició transformando los valores a minutos para luego sumar las variables y así crear la variable horas semanales de trabajo no remunerado, que incluye a las 7 actividades consultadas a los encuestados, horas semanales de trabajo doméstico no remunerado, que incorpora las actividades de la 1 a la 5 de la pregunta I1 y el indicador de horas semanales de cuidado no remunerado, que incluye a las actividades 6 y 7 del listado. El detalle, de estas actividades se encuentra en la siguiente lista:

**Tabla 22: Listado de Actividades de Trabajo No Remunerado consultados VI EME**

N°	Actividad	Tipo de trabajo
1	Cocinar, preparar alimentos para el desayuno, almuerzo o cena, recoger la mesa, lavar la loza	Trabajo Doméstico No remunerado
2	Limpiar u ordenar la casa (barrer, trapear, limpiar, tender camas, tirar basura, u otros)	
3	Lavar, tender o planchar ropa de los integrantes del hogar	
4	Realizar compras para el hogar en tiendas, mercados o supermercados (incluir tiempos de traslados)	
5	Cuidar, pasear o dar de comer a las mascotas	
6	Alimentar, bañar, vestir, acostar o mudar a otro miembro del hogar	Trabajo de Cuidado No remunerado
7	Llevar al colegio, ayudar en las tareas escolares, leer o contar cuentos a otros miembros del hogar	

Fuente: INE, VI Encuesta de Microemprendimiento, 2019

Para cada una de las variables que formaron parte de dichos indicadores se analizó su distribución de frecuencia con el fin de observar si existían valores muy altos que pudieran

influnciar fuertemente las medidas de tendencia central de cada variable. Luego de revisar los datos, se tomó la decisión de truncar los valores diarios mayores al valor del percentil 99 de la distribución en cada una de las siete variables base que forman los indicadores mencionados en el párrafo anterior, asignándole el valor del percentil 99 a todos los casos que superan este límite.

El proceso de elaboración de variables de trabajo no remunerado también contempló la creación de la variable horas semanales de trabajo total, la que se construyó sumando las horas semanales de trabajo doméstico no remunerado a las horas los datos obtenidos en c6, sobre la cantidad de días y horas semanales dedicados al negocio o actividad por cuenta propia.

**Tabla 23: medidas de distribución horas semanales destinadas al trabajo no remunerado versión original**

Sexo	N	Media	Desv. Est.	Mediana	Coef. Variación	Máximo
Hombre	4.372	11,7	10,9	8,8	0,932	170
Mujer	2.967	30,5	19,4	26,7	0,636	243,5
<b>Total</b>	<b>7.339</b>	<b>19,3</b>	<b>17,6</b>	<b>14,8</b>	<b>0,912</b>	<b>243,5</b>

Fuente: INE, VI Encuesta de Microemprendimiento, 2019

**Tabla 24: medidas de distribución horas semanales destinadas al trabajo no remunerado versión truncada**

Sexo	N	Media	Desv. Est.	Mediana	Coef. Variación	Máximo
Hombre	4.372	11,5	10,2	8,7	0,887	83
Mujer	2.967	30,2	18,1	26,6	0,599	132
<b>Total</b>	<b>7.339</b>	<b>19</b>	<b>16,7</b>	<b>14,7</b>	<b>0,879</b>	<b>132</b>

Fuente: INE, VI Encuesta de Microemprendimiento, 2019

#### 6.4. Cálculo y ajuste de ponderadores

El factor de expansión se interpreta como la cantidad de unidades en la población que representa una unidad de la muestra. La metodología de cálculo de los factores de expansión consiste en la aplicación secuencial de tres ponderadores o ajustes: el cálculo ponderador base, el ajuste por no respuesta y la calibración.

##### 6.4.1. Cálculo ponderador base

El ponderador base se define como el factor de expansión obtenido de las probabilidades de selección, sin ajustes ni correcciones, de las viviendas en la fase 1, y la selección de los microemprendedores a participar en la fase 2. En este contexto el ponderador base está compuesto por:

1. Probabilidad de que la vivienda haya sido seleccionada y contestara en la ENE en el trimestre MAM 2019.
  2. Probabilidad de seleccionar una vivienda para la EME, dado que la vivienda posee al menos un microemprendedor.
  3. Probabilidad de seleccionar un microemprendedor, dado que su vivienda fue seleccionada.
- I. Probabilidad de que la vivienda haya sido seleccionada y respondido la ENE

En la VI EME, las personas seleccionadas corresponden a un subconjunto de personas que participaron durante el proceso de encuestaje del trimestre MAM 2019 de la ENE. Por lo tanto, uno de los insumos fundamentales del ponderador base, son los factores de expansión de las viviendas de la ENE, que dan cuenta de la probabilidad de que una vivienda haya sido seleccionada y respondido la encuesta, esta variable se identifica en base a la siguiente ecuación<sup>42</sup>:

$$P_{hij}^v = \frac{1}{F_{hij}^1} \quad (1)$$

<sup>42</sup> Para mayor detalle de la construcción del Factor de expansión de la ENE revisar el apartado II.1.1 del Informe Metodológico del diseño estadístico.



## II. Probabilidad de selección de viviendas en que reside al menos un microempendedor

La selección de los microempendedores se realiza en dos etapas. Primero, se seleccionan con igual probabilidad las viviendas que contienen al menos un microempendedor, según submuestra (mes de levantamiento ENE) al interior de cada región.

Así, la probabilidad de seleccionar una vivienda que posee al menos un microempendedor viene dada por el total de viviendas seleccionadas ( $m_R^{micro}$ ) dividido por el total de viviendas disponibles ( $M_R^{micro}$ ), lo cual se detalla en la ecuación (2):

$$p_{Rj}^v = \frac{m_R^{micro}}{M_R^{micro}} \quad (2)$$

## III. Probabilidad de selección de microempendedores dado que su vivienda fue seleccionada

Una vez escogida la vivienda se seleccionan los microempendedores. La probabilidad de seleccionar al microempendedor  $k$  al interior de la vivienda  $j$ , del hogar  $l$  y rama de actividad  $m$ , perteneciente a la región  $R$ , dado que la vivienda fue seleccionada, está determinada por la división entre el número de microempendedores seleccionados ( $S_{Rjlm}^{micro}$ ) dividido por el número de microempendedores totales ( $T_{Rjlm}^{micro}$ ), como se muestra en la ecuación (3):

$$p_{Rjklm}^{micro|v} = \frac{S_{Rjlm}^{micro}}{T_{Rjlm}^{micro}} \quad (3)$$

Luego la probabilidad condicional de seleccionar el microempendedor  $k$ , en la vivienda  $j$ , de la región  $R$ , viene dada por el múltiplo entre la probabilidad de seleccionar la vivienda de un microempendedor y la probabilidad de seleccionar un microempendedor dado que su vivienda fue seleccionada, lo cual se expone en la ecuación (4):

$$p_{Rjk}^{micro} = p_{Rj}^v \cdot p_{Rjklm}^{micro|v} \quad (4)$$

## IV. Construcción del ponderador base por fase

Así, calculadas las probabilidades de selección y participación de una vivienda en la ENE en el trimestre MAM 2019 y, la probabilidad de seleccionar un microempendedor desde la EME, el ponderador base está definido por la ecuación (5):

$$F_{Rjk}^{base} = \left( \frac{1}{P_{hij}^v} \right) \cdot \left( \frac{1}{p_{Rjk}^{micro}} \right) \quad (5)$$

## V. Suavizamiento del Ponderador Base

Debido a la gran dispersión de los valores del ponderador, es necesario la incorporación de un procedimiento de suavizamiento de los factores para limitar la incidencia de unas pocas unidades muestrales en los resultados poblacionales. Dicho procedimiento consiste en truncar los factores de expansión que se ubican por encima de un umbral definido, para

luego redistribuir la diferencia resultante entre los factores asociados a las unidades que se encuentran por debajo de este umbral de manera proporcional.

En el caso del factor analizado se decidió truncar a 4 veces la media del factor base para cada rama de actividad económica reducida dando como resultado un nuevo ponderador:

$$F_{Rjk,g}^{Sr}$$

#### 6.4.2. Ajuste por no respuesta (y suavizamiento)

##### I. Ponderador ajustado por falta de respuesta

En esta etapa solo son consideradas las viviendas elegibles y que cumplen con las características de pertenecer a la población objetivo. Sin embargo, se presentan algunos casos, en los que residentes de las viviendas consideradas elegibles manifiestan su voluntad de no participar en la encuesta o que existan algunas viviendas en que no se encuentren moradores al momento del levantamiento, entre otras cosas, dando origen a la no respuesta.

Según lo anterior, surge la necesidad de aplicar un ajuste, con el objetivo de lograr que las unidades que no responden sean representadas por las que sí, previendo no introducir sesgo debido a la posibilidad de que exista relación entre la no respuesta y la variable de interés.

El método a implementar para compensar la falta de respuesta es el método de estratificación mediante “propensity score”. De acuerdo, a lo indicado por Valliant (Valliant, Dever, & Kreuter, 2013), este método consiste en modelar la probabilidad de respuesta en la VI EME como la realización de un proceso de variables latentes ( $R_i^* = x_i^T \beta + u_i$ ), buscando estimar la probabilidad de responder en la encuesta ( $P(R_i^* > \theta)$ ).

El modelo que se utiliza para realizar el ajuste por no respuesta es el modelo logístico, el cual se resume en los siguientes pasos:

1. Determinar las variables que se incluirán en el modelo de regresión logística con el cual se realizará la predicción de la probabilidad de respuesta de una persona elegible.
2. A través de este modelo, calcular la probabilidad de responder de cada una de las unidades elegibles que fueron utilizadas.
3. Ordenar las unidades de menor a mayor, según la probabilidad estimada.
4. Crear los estratos o “celdas de ajustes” donde se realizarán las correcciones de no respuesta.

Una vez creadas las celdas de ajustes<sup>43</sup>, se procede a estimar el ponderador ajustado por falta de respuesta, el cual está dado por la ecuación (6):

---

<sup>43</sup> Las celdas de ajuste son varias agrupaciones de individuos que poseen características similares asociadas a responder la encuesta. En el caso específico de la EME, se generaron 6 celdas de ajuste denominadas

$$\hat{R}_c^{NR} = \frac{\sum_{k \in S_c} F_{Rjk,g}^{Sr}}{\sum_{k \in S_{c,R}} F_{Rjk,g}^{Sr}} \quad (6)$$

Así, la expresión del ponderador ajustado por no respuesta, está dada por la ecuación (7):

$$F_{Rjk}^{NR} = F_{Rjk}^{baseSr} \cdot \hat{R}_c^{NR} \quad (7)$$

De acuerdo con la metodología antes expuesta, son 6 las celdas en las cuales se realizan los ajustes por falta de respuesta.

**Tabla 25: Total de unidades elegibles, que responde y tasa de respuesta**

Celda Ajuste	Total Elegibles	Total Responde	Tasa de Respuesta	$\hat{R}_c^{NR}$
Total	<b>8.857</b>	<b>7.808</b>	<b>88,20%</b>	<b>1,13</b>
1	1.476	921	62,40%	1,60
2	1.476	1.244	84,30%	1,19
3	1.476	1.353	91,70%	1,09
4	1.476	1.411	95,60%	1,05
5	1.476	1.436	97,30%	1,03
6	1.477	1.443	97,70%	1,02

Fuente: Elaborado por el Instituto Nacional de Estadísticas.

## II. Suavizamiento del ponderador ajustado por falta de respuesta

Obtenido el factor de viviendas ajustado por no respuesta, se evaluó la pertinencia de realizar un nuevo suavizamiento. En consecuencia, se decide realizar el suavizamiento del ponderador ajustado por falta de respuesta. Se aplica el mismo umbral de truncamiento descrito anteriormente, es decir, se decidió trincar a 4 veces la media del ponderador ajustado por falta de no respuesta para cada rama de actividad económica reducida dando como resultado un nuevo ponderador suavizado.

### 6.4.3. Calibración

En general, en todas las encuestas de hogares el ponderador final o factor de expansión es calibrado, con el objetivo de alcanzar algún stock poblacional obtenido de una fuente externa a la encuesta. En el caso de la VI EME, al no existir una fuente externa del total de microemprendedores, la población objetivo es calculada a partir de una estimación del total de microemprendedores del trimestre MAM 2019, obtenido de la ENE, actualizada al período del trabajo de campo de la VI EME.

---

“sextiles” de respuesta, mediante un modelo logístico, en que se predice la probabilidad de responder, y en base a esta probabilidad que se ordena en forma ascendente o descendente, se generan seis grupos de igual cantidad de individuos ordenados por esta probabilidad.

La Tabla 21, presenta el total de microempresarios estimado a partir de la publicación de la ENE, período MAM 2019; y según total de personas estimadas con la información levantada en MAM 2019, pero con proyecciones actualizadas a la fecha de levantamiento de la EME (en adelante  $I_{gs}$ ).

**Tabla 26. Total de microempresarios estimados a partir de la ENE – período MAM 2019**

Macrozona	Sexo	Total microempresarios	
		Factor Expansión Oficial ENE - MAM 2019	Factor Expansión Información ENE - ajustado a Junio 2019
<b>Total</b>	Hombre	1.335.972	1.340.021
	Mujer	855.104	857.519
	<b>Total</b>	<b>2.191.076</b>	<b>2.197.540</b>
<b>Norte</b>	Hombre	182.146	182.879
	Mujer	121.529	121.977
	<b>Total</b>	<b>303.674</b>	<b>304.856</b>
<b>Centro</b>	Hombre	388.891	390.367
	Mujer	228.024	228.772
	<b>Total</b>	<b>616.915</b>	<b>619.139</b>
<b>Sur</b>	Hombre	246.809	247.681
	Mujer	127.550	127.997
	<b>Total</b>	<b>374.360</b>	<b>375.678</b>
<b>Metropolitana</b>	Hombre	518.126	519.095
	Mujer	378.001	378.774
	<b>Total</b>	<b>896.126</b>	<b>897.868</b>

Elaborado por el Instituto Nacional de Estadísticas.

Finalmente, el ponderador calibrado, se le asigna a cada una de las personas entrevistadas en la EME. El procedimiento de cálculo de este ponderador se resume en tres pasos:

1. Estimar el total de microempresarios según sexo, para cada macrozona a partir de la EME 2019. Es decir, se estimó el total de microempresarios a través de la utilización del ponderador de no respuesta, tal como se muestra en la ecuación (8):

$$\hat{P}_{gs} = \sum_{j=1}^{m_g} \sum_{k=1}^{p_g} F_{Rjk}^{NR} \cdot p_{jks} \quad \begin{matrix} g = 1, 2, 3, 4 \\ s = 1, 2 \end{matrix} \quad (8)$$

2. Construir el ajuste a la población total, mediante la razón entre la estimación del total de microempresarios de acuerdo con fuentes externas (ENE), y la estimación de la encuesta obtenida en el paso (1):

$$\hat{R}_{gs} = \frac{I_{gs}}{\hat{P}_{gs}} \quad (9)$$

3. Construir el Factor de Expansión final, o Ponderador Calibrado, como el producto entre el ponderador ajustado por falta de respuesta con el ajuste a la población total, calculado en el paso (2)

$$F_{gjs}^{cal} = F_{Rjk}^{NR} \cdot \hat{R}_{gs} \quad (10)$$

Al usar el ponderador calibrado, se debe tener en consideración que éste expande al total de “microemprendedores”, de sexo  $s$  y residentes en la macrozona  $g$ , estimados a partir de la Encuesta Nacional de empleo, en el trimestre móvil MAM 2019, actualizado al crecimiento poblacional de junio 2019 - mes central de levantamiento de EME.

## 7. METODOLOGÍA DE ANÁLISIS

### 7.1. Análisis descriptivo

Se realizaron análisis descriptivos con el fin de resumir los resultados arrojados por la VI EME. Estos análisis se centraron en presentar la información recolectada en cuanto a la caracterización de los microemprendedores, la caracterización de la unidad económica, es decir, de la actividad llevada a cabo por los trabajadores por cuenta propia y los empleadores encuestados., el nivel de formalidad con el que trabajaban los informantes, la relación con el sistema financiero de estos y la capacidad de generar empleo que tienen los microemprendedores.

### 7.2. Comparación con fuentes externas

Una forma común de analizar los resultados de cualquier encuesta es compararlos con fuentes de datos externos que actúen como medio de referencia o benchmark respecto al cual comparar las proporciones o medias obtenidas. A continuación, se exponen las fuentes de información interna que se utilizaron en esta versión.

- **V EME.** Al tratarse de una encuesta que forma parte de una serie de estudios, se utilizó la versión anterior de la EME para revisar los resultados en busca de cambios significativos en el panorama general de los microemprendedores en Chile. Se revisó principalmente cómo cambiaron la cantidad total de personas que podían ser catalogadas en esta categoría, así como la proporción que representaban los trabajadores por cuenta propia sobre el total de microemprendedores, también se analizaron los cambios en las ganancias de los micronegocios entre ambas versiones. También se analizaron los cambios en la motivación emprendedora, el nivel de ganancias, la formalidad del negocio, el financiamiento principal entre otras
- **ENUT 2015.** Esta sexta versión de la EME incorporó por primera vez un módulo de trabajo no remunerado, por lo que se recurrió a la Encuesta Nacional de Uso del Tiempo de 2015 para revisar la consistencia de los resultados obtenidos. Entre los elementos que se analizaron y compararon se encuentran las horas semanales dedicadas al trabajo no remunerado como tal, al trabajo doméstico no remunerado, al trabajo de cuidado no remunerado y al trabajo total, así como las principales

diferencias entre la estimación de la ENUT y la de la EME.

- **ESI 2018.** La EME y la Encuesta Suplementaria de Ingresos (ESI) obtienen sus marcos muestrales a partir de la Encuesta Nacional de Empleo. Esta situación llevó que ambas encuestas compartieran un total de 3.854 informantes, generándose la oportunidad de comparar algunos resultados relevantes entre ambos estudios, en especial lo que tenía que ver con las ganancias de los microemprendedores. La variable analizada fue la ganancia de los informantes, revisándose la estructura de ésta y la consistencia entre las respuestas entregadas en la EME y ESI considerando las diferencias metodológicas y temporales en la captura de las preguntas de ambas mediciones.

## **8. ALCANCE DE LA PRESENTACIÓN DE RESULTADOS**

El siguiente apartado detalla los productos estadísticos que serán difundidos en el momento de la publicación, la fecha de ésta y las normas que salvaguardan el correcto uso de la información, la confidencialidad y el secreto estadístico.

### **8.1. Productos estadísticos para publicar**

Una vez finalizado el proceso de producción y análisis de la Encuesta se publican una serie de productos estadísticos a público, en la sexta versión los productos publicados se pueden dividir en: publicaciones, metodología, manuales, formularios, comités y notas técnicas, bases de datos y metadatos. Los cuales se describen a continuación:

#### **8.1.1. Publicaciones y anuarios**

Al momento de la divulgación de la base de datos, se publicó la Síntesis de la VI Encuesta de Microemprendimiento, la cual resume, por medio de un cuadernillo de 32 páginas, los principales resultados de la Encuesta abordando temas como la motivación emprendedora, el nivel de ganancias de los microemprendedores, el empleo generado y sus características, entre otras. Además, para esta versión se publicaron algunos resultados del nuevo módulo de Trabajo no remunerado.

#### **8.1.2. Metodología**

En el apartado de las metodologías, como todos los años se puede descargar el informe Metodológico del Diseño muestral. Además, para esta versión de la encuesta se encontrarán en el apartado el presente documento, el cual describe las principales metodologías implementadas en el proceso de producción de la VI Encuesta de Microemprendimiento.

#### **8.1.3. Manuales**

En relación con los datos producidos, tanto el INE como el Ministerio de Economía pone a disposición de los Usuarios en sus plataformas WEB, el Manual de Usuarios de la base de datos, el cual tiene por finalidad explicar la estructura de los datos de las bases publicadas, las variables construidas, especificación del diseño muestral, algunos análisis que pueden realizar desde la base, etc.

#### **8.1.4. Formularios**

Se puede también encontrar en formato digital el formulario de levantamiento el que fue aplicado de forma impresa a los encuestados seleccionados a participar en el estudio.

#### **8.1.5. Bases de datos**

Al igual que la versión anterior de la Encuesta se pone a disposición del público dos Bases de datos: la Base Full, que cuenta con toda la información del cuestionario salvo de la pregunta F2, y la Base de Empleo que almacena la información correspondiente a la pregunta F2 sobre la caracterización de los trabajadores del negocio. El motivo de separar

estas fuentes de información es que tienen una estructura distinta, en la base de empleo cada fila corresponde a la información de un trabajador del Microemprendimiento por lo que un encuestado puede estar presente en más de una fila dependiendo de cuantos trabajadores tenga. En cambio, en la base full cada fila corresponde a un entrevistado.

#### **8.1.6. Metadatos**

Los metadatos corresponden a información de los datos producidos describiendo los principales productos publicados. En esta sección, se encuentran disponibles las fichas técnicas de las encuestas de Microemprendimiento desde la tercera versión. Esta ficha contiene un resumen sobre las principales características de la encuesta, el tipo de muestreo, el tamaño muestral, la cobertura geográfica, entre otras.

### **8.2. Difusión de los resultados**

La difusión de la Base de datos de la Encuesta junto con toda la documentación anteriormente mencionada se realizará en marzo del 2020 por medio de los sitios web del Ministerio de Economía y el Instituto Nacional de estadísticas.

### **8.3. Accesibilidad a los datos y marco legal**

Como todas las encuestas a Hogares, la EME posee información sensible de los informantes que participaron del estudio, por ello es necesario seguir una serie de procedimientos para resguardar la identidad de los encuestados y dar cumplimiento a la normativa legal que rige la divulgación de los datos y el uso de estos mismos. A continuación, se detalla el marco legal bajo el cual se regula la liberación de los datos en la Institución y el procedimiento de anonimización realizado en la EME para dar cumplimiento a la norma.

#### **8.3.1. Marco legal**

El Marco legal que rige la divulgación de estadísticas del INE tiene dos fines: el primero, proteger la identidad de las personas y/o empresas que participan de los estudios realizados por la Institución y, el segundo, proteger la autoría de las cifras producidas por la Institución.

En relación con la Declaración de confidencialidad de los datos la ley orgánica 17.374 establece en su artículo 29 que: "El Instituto Nacional de Estadísticas, los organismos fiscales, semifiscales y Empresas del Estado, y cada uno de sus respectivos funcionarios, no podrán divulgar los hechos que se refieren a personas o entidades determinadas de que hayan tomado conocimiento en el desempeño de sus actividades. El estricto mantenimiento de estas reservas constituye el "Secreto Estadístico". Su infracción por cualquier persona sujeta a esta obligación hará incurrir en el delito previsto y penado por el artículo 247, del Código Penal, debiendo en todo caso aplicarse pena corporal."



El artículo 30 establece que: “Los datos estadísticos no podrán ser publicados o difundidos con referencia expresa a las personas o entidades a quienes directa o indirectamente se refieran, si mediare prohibición del o los afectados”. De esta manera será posible publicar datos que permitan identificar a los informantes siempre que estos tengan la venía de las personas o instituciones afectadas.

Por otra parte, de acuerdo con el derecho de autor, es necesario considerar que según la Ley N° 17.336 sobre Propiedad Intelectual, en su inciso primero del Artículo 88: “El Estado, los Municipios, las Corporaciones oficiales, las Instituciones semifiscales o autónomas y las demás personas jurídicas estatales serán titulares del derecho de autor respecto de las obras producidas por sus funcionarios en el desempeño de sus cargos”.

### 8.3.2. **Proceso Anonimización Base de datos**

El proceso de anonimización de las bases de datos implementado por el INE se basa en la metodología desarrollada por Benschop (2018) la cual consta de 12 pasos que permiten identificar el riesgo de divulgación de las bases y establecer ciertas medidas para disminuir este riesgo y enmascarar la identidad de las unidades riesgosas. A continuación, se detalla el procedimiento de anonimización desarrollado en las EME dividiéndolo en dos partes: preparación de los datos, eliminación de variables identificadoras e identificación de variables claves y/o sensibles (la cual abarca los pasos 1 al 5) e implementación del proceso de enmascaramiento (que abarca los pasos 6 al 12).

#### I. Preparación de data previo al proceso de anonimización

A continuación, se detalla la implementación del proceso de anonimización realizada en la Encuesta de Microemprendimiento (EME) dividida en tres partes: Preparación de los datos previo al proceso de anonimización (la cual abarca los pasos 1 al 5), implementación del proceso de enmascaramiento (que abarca los pasos 6 al 10) y auditoría del proceso y liberación de datos (pasos 11 y 12).

#### **Paso 1: Necesidad de tener control de la divulgación**

El objetivo del proceso de anonimización es salvaguardar la confidencialidad de la información, eliminando la posibilidad (o disminuyendo la probabilidad) de reidentificación de los informantes que participan en los estudios realizados por la institución.

La EME es una encuesta mixta que permite obtener información sensible tanto del trabajador como de su unidad económica, por lo que es necesario tener en consideración estas dos unidades de observación al momento de realizar el proceso de anonimización.

Por otra parte, si bien la EME es una Encuesta a Hogares la información publicada no corresponde a la de todo el hogar sino solo a los miembros clasificados como microemprendedores, por lo que no es posible identificar una estructura jerárquica completa

entre el hogar y el individuo. Debido a esto no se considerará un análisis del riesgo jerárquico a nivel del hogar.

## **Paso 2: Preparación de los datos y exploración de las características de los datos**

La base bruta contiene datos de 7.808 personas, de los cuales solo 7.340 mantuvieron su condición de trabajador independiente ( $a_1=1$  o 2). De ellos, 6.224 trabajaban por cuenta propia y 1.116 tenían al menos a una persona trabajando de forma asalariada. Además, contiene un total de 406 variables.

Producto de este proceso se excluyeron 14 variables, de las cuales 2 correspondían a identificadores del hogar y encuesta, 6 a variables de la hoja de ruta, 2 a variables relacionadas con la ENE y 4 a variables del proceso de imputación.

## **Paso 3: Tipo de liberación**

El tipo de liberación a analizar será una base disponible a todo público denominada como PUF (Public Use File), debido a la política institucional de transparencia y los estándares derivados de la Ley de secreto estadístico. Al estar disponible a todo público se requiere la máxima confidencialidad de la información. Esta Base se encontrará disponible en los sitios web del Ministerio de Economía y del INE.

## **Paso 4: Escenarios de intrusos y elección de variables claves**

En el caso de la EME es probable que un intruso pueda querer identificar a los informantes para conocer alguna característica particular de su negocio o su actividad independiente, como el nivel de ganancias, la formalidad del negocio, el pago de impuestos, la cantidad de trabajadores que tiene, etc.

Se analizaron las bases de datos disponibles de potenciales usuarios intrusos (Instituciones Bancarias y de servicios), quienes pueden contar con información relevante de informantes (por ejemplo) que les permitieran identificarlos en la base de datos de la encuesta a partir de una combinatoria de variables disponibles en ambas bases de datos. Ejemplo: región, nivel educacional, nacionalidad y edad.

Por ello se decidió evaluar el riesgo individual Hogares considerando la composición del Hogar a nivel individual a partir de la combinación de las siguientes variables claves: región, edad, nivel educacional, nacionalidad, rama de actividad económica, categoría ocupacional (cise) y tipo de registro ante el servicio de Impuestos Internos (SII).

## **Paso 5: Usos de la clave de datos y utilidad**

El objetivo de la EME es poder realizar una profunda caracterización de los microemprendedores y sus negocios. Una de las principales áreas de investigación de

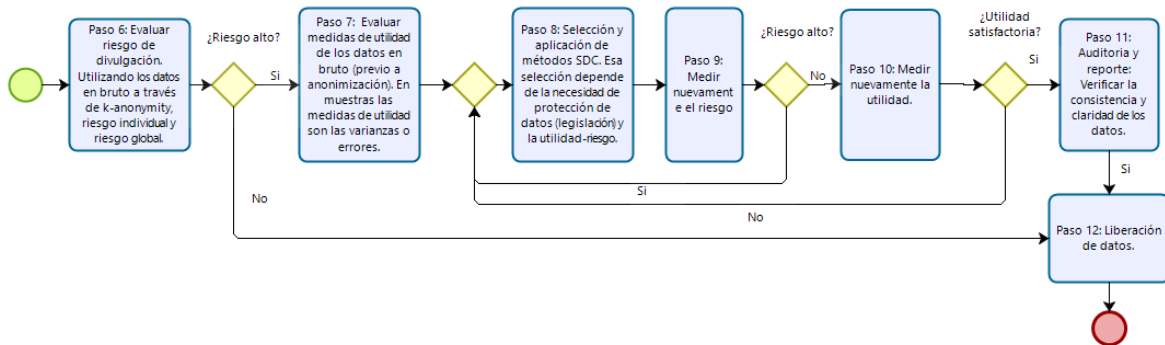
estas unidades son la informalidad, la generación de empleo, la rama de actividad económica, el acceso al sistema financiero y algunos análisis por región.

En el análisis de riesgo-utilidad de la divulgación se consideraron los cambios en la utilidad de los indicadores y estadísticas mencionados en el párrafo anterior.

## II. Implementación del proceso de enmascaramiento

El proceso de enmascaramiento de la base de datos fue realizado a nivel de individuos, sin embargo, también fueron consideradas algunas variables que entregan información de su unidad económica como la rama de actividad, el tipo de registro ante el SII y la categoría ocupacional (que permite saber si el negocio tiene trabajadores asalariados de forma permanente o no). Para ambos se implementó el procedimiento a partir de los pasos 6 al 10 y luego los pasos 11 y 12 que detallan la auditoria y liberación de datos.

**Figura 1: Proceso de enmascaramiento**



### *Paso 6b: Evaluación del riesgo (nivel individual)*

En la evaluación del riesgo individual se consideró que un intruso contaba con la información de la región, la edad, el nivel educacional, la nacionalidad, la rama de actividad económica donde trabaja, la categoría ocupacional (cise) y tipo de registro ante el SII. La siguiente tabla detalla el número de categorías de respuesta para cada una de las variables antes mencionadas:

**Tabla 27: Variables clave y número de categorías**

Variable clave	Número de categorías
<b>región</b>	16
<b>edad</b>	78
<b>cine</b>	10
<b>nacionalidad</b>	21
<b>c1_caenes_1d</b>	21
<b>cise</b>	3
<b>E3 (Tipo de registro ante el SII)</b>	8

Al evaluar el riesgo individual a partir de las 7 variables clave definidas se observa que, de las 7.808 personas, 5.301 (67,8%) incumplen el 2 anonimato, 6.852 (87,7%) el 3 anonimato y 7.557 (96,8%) el 5 anonimato, generando un riesgo de identificación bastante elevado, el cual se debe a la gran cantidad combinaciones posibles a partir de las 7 variables clave<sup>44</sup>.

Paso 7b: Evaluación de la utilidad (nivel individual)

El diseño muestral de la EME permite entregar estimaciones representativas de la variable de interés (Cuenta propia /Total de microemprendedores) para las 16 regiones del país. No obstante, no es posible asegurar confiabilidad en las estimaciones a un mayor nivel de desagregación o cruces de variables debido a la poca cantidad de casos muestrales.

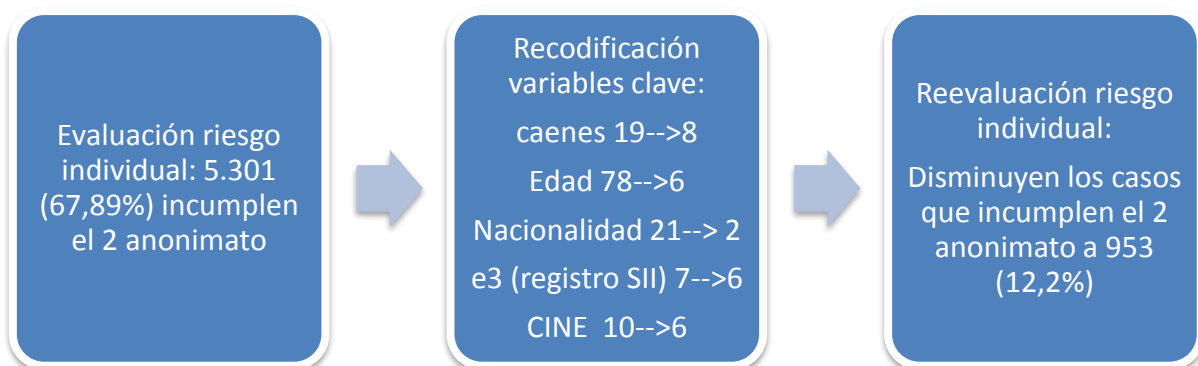
Paso 8b: Elección y aplicación de métodos SDC (nivel individual)

Se evaluaron distintas medidas para disminuir el riesgo, optando por aplicar tan sólo el método de recodificación de variables. La implementación de este proceso consideró también la evaluación de la supresión local, la que fue descartada debido al impacto generado en el indicador principal de la encuesta a nivel regional, donde el proceso de anonimización aumentaba la participación de los Microemprendedores por cuenta propia en todas las regiones. El proceso de anonimización final consideró los siguientes pasos:

---

<sup>44</sup> Es importante notar que tan sólo con las variables de edad, nacionalidad y rama de actividad un intruso puede generar un número de combinaciones mayor a 100.000 ( $78 \times 21 \times 68 = 111.384$ ) por lo que el posible riesgo es bastante alto.

## Diagrama 7: Implementación métodos SDC y evaluación del riesgo



Dado el alto riesgo de identificación de variables se recodificó la variable de rama de actividad económica, disminuyendo el número de categorías de 21 a 8. En cuanto a la variable de edad se recodificó la variable continua en una categórica, pasando 106 posibles respuestas a 6 distribuidas en los siguientes tramos, 15-24, 25-34, 35-44, 45-54, 55-64, 65 o más. La variable de registro ante el SII se recodificó, fusionando las categorías 3 “Sí, como Empresa individual...” y 4 “Sí, como otro tipo de empresa”. La variable de CINE consideró la agrupación de las categorías “Nunca Estudió” y “Educación preescolar” y Nivel ignorado en “Nunca estudió”, las categorías de “Educación primaria (nivel 1)” y Educación primaria (nivel 2) en “Educación primaria” y las categorías “Postítulo y maestría” y “Doctorado” en “Postítulo, Magíster o Doctorado” reduciendo a 6 las categorías de respuesta. Finalmente, en el caso de la nacionalidad se decidió reducir a dos las categorías 0 “nacional” y 1 “extranjero”.

### Paso 9b y 10b: Reevaluación del riesgo (nivel individual)

Al aplicar los métodos SDC, a nivel individual se observa que se logró reducir el riesgo.

En cuanto al impacto en la utilidad de la recodificación de las variables, el proceso de anonimización, no disminuye la utilidad de la base debido a la poca confiabilidad que entregaban las desagregaciones a mayor nivel de los datos. De hecho, a nivel institucional ninguna de las recodificaciones afecta los tabulados y publicaciones realizados por lo que no existen inconvenientes en publicar la base anonimizada.

### III. Auditoría del proceso y liberación de datos

A partir de los resultados de este ejercicio, dado el alto riesgo de identificación de informantes la base de datos publicada contiene las variables de edad, caenes, tipo de registro ante el SII, y cine recodificadas, según lo detallado en el paso 8.

## 9. GLOSARIO

**Actividad económica:** es una combinación de acciones cuyo resultado es la producción de bienes o servicios.

**Actividad no observada:** actividad comprendida en el ámbito de la producción del Sistema de Cuentas Nacionales que no ha sido observada ni medida a través de los datos básicos que sirven para la elaboración de las Cuentas Nacionales.

**Asalariado/ Trabajador dependiente:** se refiere a aquellas personas que cuentan con un contrato de trabajo, ya sea implícito o explícito, por el que reciben una remuneración. Pueden ser clasificados como asalariados del sector privado, o bien, del sector público.

- **Asalariado del sector privado:** corresponden a aquellas personas que trabajan para un empleador privado y perciben un sueldo como remuneración (mensual, semanal, quincenal, etc.) Se excluyen de esta categoría, aquellas personas que trabajan como servicio doméstico puertas adentro, o puertas afuera.
- **Asalariado del sector público:** se refiere a aquellas personas que trabajan para un organismo público, ya sea gobierno central o municipalidades, FF.AA. o empresas públicas; es decir, aquellas de propiedad mixta o estatal (Ej. Correos de Chile, Metro, Banco Estado, entre otras) y que reciben un sueldo o remuneración (mensual, quincenal, semanal, etc.). Son parte de este grupo también, aquellas personas que trabajen en organismos o institucionales centralizadas, como Ministerios, Intendencias, Gobernaciones, Servicios Públicos, entidades estatales autónomas, Poder Judicial, entre otros, así como también, quienes trabajan en instituciones públicas sin fines de lucro, como establecimientos educacionales (Universidades e Institutos Profesionales) con aporte estatal y de salud municipalizados.
- **Servicio Doméstico:** son todas aquellas personas que trabajan para un empleador o patrón y se dedican a actividades propias del servicio del hogar (aseo, preparación de alimentos, lavados, etc.) Esta categoría incluye a asesoras del hogar, jardineros, choferes, entre otros, quienes pueden residir o no, en la vivienda seleccionada. Debemos tener presente que para clasificar a un informante en esta categoría, **debe mantener contrato con un solo empleador, ya que si tiene más de un empleador, debe ser clasificado como trabajador por cuenta propia.**

**Aprendices:** se refiere a aquellas personas asalariadas que poseen un contrato donde se especifica que toda, o una parte de su remuneración, consiste en capacitación para el ejercicio de un oficio o profesión.

**Banco Central:** el banco central es la institución financiera nacional que ejerce el control sobre los aspectos claves del sistema financiero.

**Bases de datos:** consisten en archivos de datos organizados de tal forma que permitan el acceso a los datos y su utilización Efectiva.

**Categoría ocupacional:** establece la relación de dependencia y el riesgo que asume un trabajador y cómo este se vincula con el mercado. Estas categorías son: trabajador por cuenta propia (independiente o empleador), trabajador dependiente o asalariado (del sector privado, público o de los hogares) y trabajador familiar no remunerado.

**Clasificación Industrial Internacional Uniforme (CIIU):** la CIIU es la Clasificación Industrial Internacional Uniforme, por industria, de todas las ramas de actividad económica de las Naciones Unidas. Esta clasificación es la norma internacional de clasificación de las actividades económicas productivas. Su principal objeto es el de proporcionar un conjunto de categorías de actividad de modo que las entidades puedan clasificarse según su actividad económica. La organización jerárquica de la clasificación comprende los siguientes niveles: a) Divisiones - códigos con dos dígitos; b) Grupos – códigos con tres dígitos; c) Clases – códigos con cuatro dígitos.

**Bienes:** los bienes son objetos físicos los cuales tienen una demanda. Si existe una demanda, significa que los bienes pueden utilizarse para cubrir las necesidades de los hogares o de la colectividad, o para producir otros bienes u otros servicios.

**Bien durable:** un bien durable es aquel que puede utilizarse repetida o continuamente durante un período superior a un año, asumiendo una tasa normal o promedio de uso físico.

**Bienes privados:** los bienes privados los cuales tienen derechos de propiedad, la cual puede ser transferida de una unidad institucional a otra, a través de una operación en el mercado. Si existe una demanda, significa que los bienes pueden utilizarse para cubrir las necesidades de los hogares o de la colectividad, o para producir otros bienes u otros servicios.

**Comisión Económica para América latina y el caribe (CEPAL):** es una de las cinco comisiones regionales de las Naciones Unidas y su sede está en Santiago de Chile. Se fundó para contribuir al desarrollo económico de América Latina, coordinar las acciones encaminadas a su promoción y reforzar las relaciones económicas de los países entre sí y con las demás naciones del mundo. Posteriormente, su labor se amplió a los países del Caribe y se incorporó el objetivo de promover el desarrollo social.

**Consumo:** consiste en la utilización de bienes y servicios para la satisfacción de necesidades o deseos humanos individuales o colectivos

**Consumo Final de los Hogares:** comprende los gastos que realizan los hogares en la compra de bienes duraderos y no duraderos y servicios para la satisfacción directa de sus necesidades. Se excluye los gastos que realizan las familias en la adquisición de viviendas

considerándose estos, como formación bruta de capital de la economía, igualmente se excluyen las compras de objetos de valor que no se deterioran con el paso del tiempo tales como las obras de arte, piedras y metales preciosos.

**Consumo intermedio (CI):** consiste en el valor de los bienes y servicios consumidos como insumos por un proceso de producción, excluidos los activos fijos cuyo consumo se registra como consumo de capital fijo.

**Constitución legal:** corresponde a dos alternativas de inscripción en el SII: persona natural o persona jurídica. La diferencia entre ambas es que en el caso de la persona jurídica es la empresa la que asume las obligaciones y los derechos de la misma, mientras que, en la constitución como persona natural, es el dueño quien asume dichas responsabilidades.

**Contabilidad completa y formal:** registro y resumen periódico de todas las transacciones económicas de la unidad económica. Permite conocer la Situación Patrimonial y el Estado de pérdidas y ganancias de la unidad económica al finalizar un período contable. Según el Diccionario Tributario del SII, “comprende los libros Caja, Diario, Mayor e Inventarios y Balances, independiente de los libros auxiliares que exija la ley, tales como Libro de Ventas Diarias, de Remuneraciones, de Impuestos Retenidos, etc.”

**Cuasisociedad:** unidad económica dedicada a producir bienes y servicios para el mercado. No cuenta con reconocimiento legal como persona jurídica y puede ser de propiedad de uno o más miembros del hogar. Cuentan con un sistema contable que permite separar gastos del negocio del hogar, comportándose de alguna manera como una unidad económica formal.

**Desocupados:** todas las personas en edad de trabajar que no estaban ocupadas durante la semana de referencia (entre lunes y domingo), que habían llevado a cabo actividades de búsqueda de un puesto de trabajo durante las últimas cuatro semanas (incluyendo la de referencia) y que estaban disponibles para trabajar en las próximas dos semanas (posteriores a la de referencia).

**Economía informal:** conjunto de “todas las actividades económicas de los trabajadores y las unidades económicas que están —en la legislación o en la práctica— no cubiertas o insuficientemente contempladas por sistemas formales” Fuente: 90ª Reunión de la Conferencia Internacional del Trabajo (2002). Este concepto se compone de dos dimensiones: Sector Informal y Ocupación Informal.

**Economía no observada:** corresponden a las actividades subterráneas, ilegales o informales o aquéllas que son realizadas por los hogares para consumo final propio. También puede haber actividades no observadas debido a vacíos en el dispositivo de captura de datos estadísticos básicos.

**Economía subterránea:** son actividades que pueden ser productivas desde el punto de vista económico y totalmente legales (a condición de que cumplan con algunas normas o



normativas), pero deliberadamente ocultadas a las autoridades públicas por distintos tipos de razones: a) evitar el pago del impuesto a la renta, del impuesto sobre el valor agregado o de otros impuestos; b) evitar el pago de las cotizaciones del seguro social; c) no tener que cumplir con algunas normas legales: salario mínimo, horario máximo, normas en materia de seguridad o de salud, etc.; d) evadir algunos procedimientos administrativos, tales como el hecho de llenar cuestionarios estadísticos u otros formularios administrativos.

**Economía total:** se define como el conjunto de todas las unidades institucionales residentes en el territorio económico de un país.

**Empleador:** persona que dirige su propia empresa y que contrata de forma permanente los servicios de uno o más trabajadores a cambio de una remuneración.

**Empresa/Unidad económica:** toda unidad de producción de bienes o servicios para la venta o el trueque. Comprende tanto a las unidades que contratan trabajadores como a aquellas que pertenecen o son administradas por un trabajador por cuenta propia. Las actividades de producción pueden realizarse dentro o fuera de un establecimiento, así como también en la calle, con o sin un lugar fijo.

**Emprendedores:** son trabajadores que crean trabajo para sí mismos y potencialmente para otros por medio del establecimiento y operación de una empresa. La categoría de “trabajadores independientes” de la clasificación de la situación en la ocupación es la más adecuada para identificar a los emprendedores y recopilar estadísticas al respecto.

**Establecimiento:** un establecimiento se define como una empresa o como parte de una empresa que ejerce un solo tipo de actividad económica, de manera independiente, exclusiva o principal, en un solo lugar o a partir de un solo lugar, o dentro de una zona geográfica, y para la cual están disponibles datos que permiten calcular el excedente de explotación o que pueden ser compilados de manera significativa.

**Excedente de explotación:** se refiere a la retribución al riesgo empresarial y constituye por lo tanto las ganancias o pérdidas derivadas de una actividad productiva. Para las empresas constituidas en sociedad se diferencia de las utilidades en que éstas últimas incluyen resultados de ingresos que no se derivan de actividades productivas. Para las empresas no constituidas en sociedad, específicamente las empresas individuales, trabajadores por cuenta propia o trabajadores independientes que en el proceso de producción utilizan factores de producción como trabajadores familiares no remunerados y activos fijos que pueden ser utilizados para la producción y a la vez para el consumo final, el excedente de explotación se denomina “ingreso mixto”, debido a que no es posible distinguir del excedente propiamente dicho las remuneraciones y el consumo de capital fijo.

**Frontera de la producción:** en el Sistema de Cuentas Nacionales la frontera de la producción incluye las siguientes actividades: (a) la producción de todos los bienes o servicios que se suministran, o que se pretende suministrar, a unidades distintas de

aquéllas que los producen, incluida la producción de los bienes o servicios utilizados completamente en el proceso de producción de dichos bienes o servicios; (b) la producción por cuenta propia de todos los bienes que sus productores conservan para su autoconsumo final o para su formación bruta de capital; (c) La producción por cuenta propia de productos de captura del conocimiento que sus productores conservan para su autoconsumo final o formación bruta de capital, pero con exclusión (por convención) de este tipo de productos cuando son producidos por los hogares para su propio uso; (d) la producción por cuenta propia de los servicios de las viviendas ocupadas por sus propietarios; y (e) la producción de los servicios domésticos y personales producidos por personal de servicio doméstico remunerado.

**Hogar:** un hogar es un grupo de personas que comparten la misma vivienda, que juntan, total o parcialmente, su ingreso y su riqueza y que consumen colectivamente ciertos tipos de bienes y servicios, sobre todo los relativos a la alimentación y alojamiento.

**Fuerza de trabajo:** personas en edad de trabajar que durante la semana de referencia cumplen los requisitos para ser incluidas en la categoría de ocupados o desocupados. Horas habituales: corresponde a las horas comúnmente trabajadas en un período de referencia determinado (promedio de los últimos tres meses).

**Instituciones sin fines de lucro (ISFL):** “entidades jurídicas o sociales creadas para producir bienes y servicios, cuyo estatuto jurídico no les permite ser fuente de ingreso, beneficios u otras ganancias financieras para las unidades que las establecen, controlan o financian” (SCN, 2008, p. 72).

**Ingreso Mixto:** el saldo de la cuenta de generación del ingreso de las empresas no constituidas en sociedades que pertenecen al sector de los hogares se llama ingreso mixto. En efecto, en este caso, el saldo de la cuenta de generación del ingreso contiene implícitamente un elemento de remuneración para el trabajo efectuado por el propietario o los miembros de su familia que no puede distinguirse de su beneficio como empresario.

**Microempendedor:** empleador o trabajador por cuenta propia dueño de una microempresa (unidad económica con hasta 10 trabajadores, incluyendo al informante que es dueño de la unidad económica).

**Microempresa:** unidad económica de hasta 10 trabajadores incluido el dueño del negocio.

**Microempresa familiar:** empresa que pertenece a una o a más personas naturales que realizan sus actividades empresariales en la casa habitación en que reside uno de ellos. • Régimen de microempresa familiar (MEF): procedimiento simplificado de iniciación de actividades para microempresas familiares.

**Ocupados:** todas las personas en edad de trabajar, que durante la semana de referencia dedicaron al menos una hora a alguna actividad para producir bienes o servicios a cambio de una remuneración o beneficios.

**Ocupación informal:** se considera que tienen una ocupación informal aquellos dependientes que no cuentan con cotizaciones de salud y previsión social por concepto de su vínculo laboral con un empleador. Para el caso de los empleadores y trabajadores por cuenta propia, se estima que poseen una ocupación informal si la empresa, negocio o actividad que desarrollan pertenece al sector informal. En el caso de familiar no remunerado (del hogar), por definición todos tienen una ocupación informal dada las condiciones de su vinculación con la unidad económica donde se desempeña.

**Ocupación Formal:** tienen una ocupación informal todos aquellos ocupados dependientes que cuentan con cotizaciones de salud (Isapre o Fonasa) y previsión social (AFP u otro sistema de previsión) por concepto de su vínculo o relación laboral con un Empleador, teniendo que cumplirse ambos criterios de forma simultánea para ser clasificados con una Ocupación Formal. Para el caso de los Empleadores y Trabajadores por Cuenta Propia, se considera que poseen una Ocupación Formal si la empresa, negocio o actividad que desarrollan pertenece al Sector Formal.

**Ocupación en el Sector Informal:** conjunto de Ocupados cuyo trabajo principal se realiza dentro de unidades económicas pertenecientes al Sector Informal de la economía. Esto considera personas tanto con ocupación informal como formal.

**Organización Internacional del Trabajo (OIT):** única agencia 'tripartita' de la Organización de las Naciones Unidas (ONU), que reúne a gobiernos, empleadores y trabajadores de 187 Estados miembros a fin de establecer las normas del trabajo, formular políticas y elaborar programas promoviendo el trabajo decente de todas y todos, mujeres y hombres.

**Organización para la Cooperación y el Desarrollo Económico (OCDE):** agrupa a 35 países miembros y su misión es promover políticas que mejoren el bienestar económico y social de las personas alrededor del mundo. Ofrece un foro donde los gobiernos puedan trabajar conjuntamente para compartir experiencias y buscar soluciones a los problemas comunes.

**Pensión básica solidaria de invalidez (PBSI):** corresponde a un beneficio económico mensual entregado por el Estado a personas entre 18 y 65 años, declaradas con invalidez y que no tienen derecho a pensión en ningún régimen previsional. Las personas beneficiarias recibirán esta pensión hasta el último día del mes en que cumplan 65 años. La Pensión Básica Solidaria de Invalidez es incompatible con cualquier otra pensión de algún régimen previsional.

**Pensión básica solidaria de vejez (PBSV):** beneficio económico mensual entregado por el Estado a las personas de 65 o más años, que no tienen derecho a pensión en algún

régimen previsional, y que integren un grupo familiar perteneciente al 60% más pobre de la población de Chile.

**Población en edad de trabajar (PET):** población actualmente residente en el país, de 15 años y más. Registro en Servicio de Impuestos Internos (SII): corresponde a la formalización de una unidad económica ante el Servicio de Impuestos Internos. El SII es una institución pública chilena dependiente del Ministerio de Hacienda, encargada, especialmente, de la aplicación y fiscalización administrativa de las disposiciones tributarias.

**Producción ilegal:** producción de bienes y servicios cuya venta, distribución o posesión está prohibida por la ley; actividades de producción que generalmente son legales pero que se vuelven ilegales cuando las realizan productores no autorizados. Ejemplos: tráfico de drogas y el tratamiento médico ilegal.

**Producción oculta (subterránea):** actividades de producción que son legales cuando se realizan en cumplimiento de las normas o reglamentos pertinentes, pero que se ocultan deliberadamente a las autoridades públicas para: (1) evitar el pago del impuesto a la renta, (2) evitar el IVA u otros impuestos, (3) evitar el pago de las 55 contribuciones de la seguridad social, y (4) evitar el cumplimiento de las normas legales tales como salarios mínimos, el límite legal de horas de trabajo, las normas de seguridad y salud, etc.

**Registro Contable:** contabilidad sistemática y ordenada de los ingresos y egresos de una unidad económica, los que deben reflejar las operaciones y transacciones realizadas por el negocio en un período de tiempo determinado. En algunas circunstancias estos registros pueden ser objetivo de monitoreo o auditoría por parte de alguna autoridad tributaria.

**Régimen de tributación simplificada:** forma de tributación para micro y pequeños empresarios mediante una contabilidad simplificada, que los libera parcialmente de la contabilidad completa a través de la deducción inmediata de gastos de las inversiones e inventarios, entre otros. Pueden acceder a este régimen empresarios individuales o de una Empresa Individual de Responsabilidad Limitada (EIRL) que califiquen como contribuyente en primera categoría y que cumplan los siguientes requisitos: (1) mantener vigente el inicio de actividades, (2) sean contribuyente de IVA, (3) tener por giro de actividad rentas inmobiliarias, capitales de renta mobiliarios, salvo las necesarias para el desarrollo de su actividad principal, y (4) que el promedio anual de sus ingresos de giro no supere las 5.000 Unidades Tributarias Mensuales (UTM) en los últimos tres ejercicios.

**Régimen simplificado de impuesto a la renta:** es un beneficio tributario al cual pueden acogerse todos los contribuyentes obligados a llevar una contabilidad completa y que no hayan excedido un promedio anual de 3.000 unidades tributarias mensuales (\$557.532 abril 2017). Este régimen da la opción de tributar sobre el dinero retirado y no sobre las ganancias del negocio, como se realiza de forma habitual.

**Régimen simplificado de IVA:** refiere a un régimen de tributación para pequeños comerciantes, artesanos y pequeños prestadores de servicios, que venden o prestan servicios directamente al público o consumidor. Su principal característica es liberarlos de la obligación de emitir boletas y cumplir con el Débito Fiscal (pago del IVA), lo cual debe ser pactado con anterioridad con el SII.

**Sector Informal:** lo constituyen todas aquellas unidades económicas de mercado que no cuenten con registro en el Servicio de Impuestos Internos (SII) ni tampoco puedan ser clasificadas como cuasi-sociedades, ya que no poseen una contabilidad completa o simplificada que les permite realizar una efectiva separación de sus gastos. Se excluyen de este sector a las unidades de no mercado como son; sector gobierno, instituciones sin fines de lucro, organismos internacionales, según lo define el Sistema de Cuentas Nacionales del año 2008. Además, se excluyen los Hogares como Empleadores que contratan los servicios de los trabajadores domésticos.

**Sector Formal:** son todas las unidades económicas que cuentan con registro en el Servicio de Impuestos internos (SII) o en su defecto, aquellas unidades económicas que carecen de registro, pero pueden ser clasificadas como cuasi-sociedades. Se incluyen en el grupo de las cuasi- sociedades a aquellas unidades económicas que reconocen tener contabilidad simplificada, ya que, independiente de que no cumplan las condiciones estrictas de una contabilidad completa, responden a una política pública creada para incentivar la formalización de las pequeñas unidades económicas. De forma adicional, y en línea con las recomendaciones internacionales, se incluye en el Sector Formal, a las unidades no de mercado (Sector Público, Instituciones sin fines de lucro y Organismos internacionales) según lo define el Sistema de Cuentas Nacionales del año 2008

**Sector Hogares como Empleadores:** comprende las actividades realizadas por Hogares que contratan los servicios de Trabajadores Domésticos remunerados, así como otro tipo de servicios al hogar. El producto resultante de esta actividad es consumido por el propio hogar, por tanto, no es una actividad orientada al mercado, quedando fuera del concepto de Sector Informal o Formal.

**Sector del gobierno general:** “comprende principalmente a las unidades gubernamentales centrales, estatales y locales junto con los fondos de seguridad social impuestos y controlados por el gobierno. Adicionalmente, incluye a las ISFL dedicadas a la producción de bienes y servicios no de mercado y controladas por el gobierno o por fondos de seguridad social” (SCN, 2008, p. 76).

**Seguro de desempleo o cesantía:** es un beneficio monetario entregado por un período de tiempo a los trabajadores regidos por el Código del Trabajo, que hayan perdido su empleo. Es entregado por la Administradora de Fondos de Cesantía (AFC) u otras instituciones privadas.

**Servicio de Impuestos Internos (SII):** es una de las instituciones fiscalizadoras del Estado. De acuerdo con la ley, sus funciones son la aplicación y fiscalización de todos los impuestos internos actualmente establecidos (o que se establecieron), fiscales o de otro carácter; en que tenga interés el Fisco y cuyo control no esté especialmente encomendado por la ley a una autoridad diferente. Existen distintos tipos de registros en el SII:

**a. Trabajador honorario (boleta de honorarios):** persona que presta servicios y percibe ingresos mediante la realización de trabajos personales. Para ello se emite un documento por los ingresos generados en las prestaciones de servicios personales, por las actividades clasificadas en la Segunda Categoría de la Ley de la Renta (no afectos a IVA).

**b. Empresa constituida como Persona natural (El RUT del negocio es igual al RUT de su dueño):** son aquellas empresas en que el dueño ejerce todos los derechos y obligaciones del negocio o actividad a su nombre. La empresa funciona con el mismo RUT de la persona natural, quien es responsable legal del patrimonio de la empresa. Puede funcionar como Empresa Individual o Microempresa Familiar.

**c. Empresa Individual de Responsabilidad Limitada (EIRL):** es una “Persona jurídica formada exclusivamente por una persona natural, con patrimonio propio y distinto al del titular, que realiza actividades de carácter netamente comercial y están sometidas a las normas del Código de Comercio, cualquiera sea su objeto, y pudiendo realizar toda clase de operaciones civiles y comerciales, excepto las reservadas por la ley a las Sociedades Anónimas” (art. 2° Ley N° 19.857).

**d. Sociedad Anónima (S.A):** “es una persona jurídica formada por la reunión de un fondo en común suministrado por accionistas responsables solo por sus respectivos aportes y administrada por un directorio integrado por miembros esencialmente revocables” (Art1. Ley N° 18.046).

**e. Sociedad por Acciones (SPA):** “la sociedad por acciones, o simplemente, es una persona jurídica creada por una o más personas mediante un acto de constitución perfeccionado de acuerdo con los preceptos siguientes, cuya participación en el capital es representada por acciones” (Código de Comercio de la República de Chile, Ob. Cit., Art. 424).

**Sistema de Cuentas Nacionales (SCN):** conjunto de recomendaciones aceptadas internacionalmente, para la elaboración de “mediciones de la actividad económica de acuerdo con convenciones contables estrictas, basadas en principios económicos” (Sistema de Cuentas Nacionales (SCN), 2008, p. 1).

**Subsidio único familiar (SUF):** corresponde a un beneficio otorgado por el Estado, a aquellas personas de escasos recursos, que no pueden acceder al beneficio de Asignación familiar o Maternal, dado que no son trabajadores afiliados a un sistema previsional.

**Sociedad:** unidad económica dedicada a producir bienes y servicios para el mercado, que cuenta con reconocimiento legal como persona jurídica (constituida) independiente de sus propietarios.

**Tasa de desocupación:** número de personas desocupadas expresado como porcentaje de la fuerza de trabajo. **Tasa de ocupación:** número de personas ocupadas como porcentaje de la población en edad de trabajar. **Tasa de ocupación en el sector informal (TOSI):** número de ocupados en unidades económicas informales expresado como porcentaje del total de ocupados.

**Tasa de ocupación informal (TOI):** número de personas con ocupaciones informales expresado como porcentaje del total de ocupados.

**Tasa de ocupación informal no agropecuaria (TOI NO AGRO):** número de personas con ocupaciones informales fuera del sector agropecuario (agricultura, ganadería, silvicultura y pesca), expresado como porcentaje del total de ocupados, excluyendo a aquellos que son dueños o cuyo sueldo es pagado por una unidad económica perteneciente al sector agropecuario.

**Tasa combinada de desocupación y tiempo parcial involuntario:** número de personas clasificadas como desocupados, iniciadores disponibles y ocupados a tiempo parcial involuntario, expresado como porcentaje de la fuerza de trabajo y los iniciadores disponibles.

**Tasa de participación:** número de personas en la fuerza de trabajo expresado como porcentaje de la población en edad de trabajar.

**Trabajo:** comprende todas las actividades realizadas por personas de cualquier sexo o edad, con el fin de producir bienes o prestar servicios para el consumo de terceros, de un familiar o para uso final propio.

**Trabajo en la ocupación:** corresponde al concepto tradicional de trabajo, comprende actividades que tienen por objetivo producir bienes o servicios a cambio de remuneración o beneficios.

**Trabajo en formación remunerado:** comprende actividades no remuneradas para la producción de bienes o servicios para terceros, con el objetivo de adquirir experiencia en el lugar de trabajo o competencias en un oficio o profesión

**Trabajo en la producción para el autoconsumo:** comprende actividades que tienen como fin producir bienes y servicios para uso final propio o del hogar.

- **Trabajo en la producción de servicios para el autoconsumo:** los cuales abarcan la contabilidad y administración del hogar (compra de bienes, transporte, etc.), preparación y servicio de comidas, limpieza, decoración y mantención de la vivienda y el cuidado y educación de niños u otras personas del hogar.
- **Trabajo en la producción de bienes para el autoconsumo:** los cuales abarcan la contabilidad y administración del hogar (compra de bienes, transporte, etc.), preparación y servicio de comidas, limpieza, decoración y mantención de la vivienda y el cuidado y educación de niños u otras personas del hogar.

**Trabajador por cuenta propia:** trabajador independiente sin relaciones de subordinación a un jefe, que ejerce por su cuenta una profesión, oficio o negocio. Trabaja solo, sin embargo, puede estar asociado con familiares, socios o trabajadores que emplea de forma ocasional o sin pago en dinero.

**Trabajador familiar no remunerado:** persona que presta apoyo en la empresa o negocio de un familiar sin recibir un pago en dinero o especies por su trabajo.

**Trabajadores independientes:** son todos aquellos trabajadores que carecen de una relación de subordinación y dependencia con un tercero. Asumen riesgos y tienen control sobre las decisiones de la unidad económica.

**Unidad de producción:** corresponde a todas aquellas entidades o unidades económicas dedicadas a la producción de bienes o servicios. Típicamente se refiere a empresas, pero esta definición incluye a organismos estatales, hogares con al menos una parte de la producción destinada al mercado e instituciones sin fines de lucro.

**Unidades Institucionales residentes:** las unidades que configuran la economía total se agrupan en cinco sectores institucionales mutuamente excluyentes. El territorio económico, aunque se ajusta básicamente al territorio geográfico, no coincide exactamente con él, efectuándose algunas adiciones y sustracciones al mismo. En el Sistema de Cuentas Nacionales, el concepto de residencia no se basa en la nacionalidad ni en criterios jurídicos. Se dice que una unidad institucional es residente de un país cuando tiene un centro de interés económico predominante en el territorio económico de dicho país; es decir, cuando realiza en él actividades económicas durante un período de tiempo prolongado (un año o más es el período adoptado como guía práctica).



## 10. BIBLIOGRAFÍA

Cochran, W. (1998). Técnicas de Muestreo. México D.F.: Compañía Editorial Continental, S.A. de C.V.

CIT (Conferencia General de la Organización Internacional del Trabajo), 90ª reunión, 2002. Resolución relativa al trabajo decente y la economía informal.

DANE Colombia (2009). Encuesta de Microestablecimientos. Metodología Encuesta de Microestablecimientos. Disponible en <https://www.dane.gov.co/files/investigaciones/fichas/Microestablecimientos.pdf>

Hart, Keith, "Informal Income Opportunities and Urban Employment in Ghana", artículo presentado en la conferencia "Desempleo urbano en África", en el Instituto de Estudios del Desarrollo (IDS) de la Universidad de Sussex, septiembre de 1971.

IBGE (Instituto Brasileiro de Geografia e Estatística). 2006. Economía informal urbana. Serie Relatorios Metodológicos, Río de Janeiro. Disponible en <http://www.ibge.gov.br/home/estatistica/economia/ecinf/2003/ECINFmetodpesquisa.pdf>.

INE Chile (2016), CAENES Clasificador de Actividades Económicas Nacional para Encuestas Sociodemográficas. Recuperado de: <http://www.ine.cl/estadisticas/laborales/ene>

INE Chile (2018), Manual conceptual y metodológico, diseño muestral, Encuesta Nacional de Empleo (ENE). Santiago. Recuperado de: [https://www.ine.cl/docs/defaultsource/laborales/ene/publicaciones/clasificador-de-actividadesecon%C3%B3micas-nacional-para-encuestas-sociodemogr%C3%A1ficas\(caenes\).pdf](https://www.ine.cl/docs/defaultsource/laborales/ene/publicaciones/clasificador-de-actividadesecon%C3%B3micas-nacional-para-encuestas-sociodemogr%C3%A1ficas(caenes).pdf)

INE Chile. (2018). Encuesta Suplementaria de Ingresos (ESI) Manual Conceptual y Metodológico. Instituto Nacional de Estadísticas de Chile. Disponible en [https://www.ine.cl/docs/default-source/encuesta-suplementaria-de-ingrseos/metodologia/documento-metodol%C3%B3gico/documento\\_metodologico\\_esi.pdf?sfvrsn=a5e286b4\\_8](https://www.ine.cl/docs/default-source/encuesta-suplementaria-de-ingrseos/metodologia/documento-metodol%C3%B3gico/documento_metodologico_esi.pdf?sfvrsn=a5e286b4_8)

INE Chile. (2018). Estadísticas de Informalidad laboral: Marco conceptual y manual metodológico. Instituto Nacional de Estadísticas de Chile. Disponible en [https://www.ine.cl/docs/default-source/informalidad-y-condiciones-laborales/metodologia/antecedentes-metodologicos/manual-conceptual-y-metodol%C3%B3gico-informalidad-laboral.pdf?sfvrsn=afad6bfc\\_5](https://www.ine.cl/docs/default-source/informalidad-y-condiciones-laborales/metodologia/antecedentes-metodologicos/manual-conceptual-y-metodol%C3%B3gico-informalidad-laboral.pdf?sfvrsn=afad6bfc_5)

INE Chile. (2018). VIII Encuesta de presupuestos Familiares (EPF) Metodología. Instituto Nacional de Estadísticas de Chile. Disponible en [https://www.ine.cl/docs/default-source/encuesta-de-presupuestos-familiares/metodologia/viii-epf---\(julio-2017---junio-2017\)/metodolog%C3%ADa-viii-epf.pdf?sfvrsn=f098409c\\_2](https://www.ine.cl/docs/default-source/encuesta-de-presupuestos-familiares/metodologia/viii-epf---(julio-2017---junio-2017)/metodolog%C3%ADa-viii-epf.pdf?sfvrsn=f098409c_2)

INEC Costa Rica (2019). Encuesta Nacional de Microempresas de los Hogares 2019. Aspectos Metodológicos. Disponible en [http://inec.cr/sites/default/files/documentos-biblioteca-virtual/meeconomenameh2019-06\\_0.pdf](http://inec.cr/sites/default/files/documentos-biblioteca-virtual/meeconomenameh2019-06_0.pdf)

INEGI México (2012). Encuesta Nacional de Micronegocios. Metodología. Disponible en [https://www.inegi.org.mx/contenidos/programas/enamin/2012/doc/enamin12\\_met\\_dis\\_mu\\_estrал.pdf](https://www.inegi.org.mx/contenidos/programas/enamin/2012/doc/enamin12_met_dis_mu_estrал.pdf)

INEI Perú (2017). Producción y Empleo Informal en el Perú. Cuenta Satélite de Economía Informal 2007-2012. Disponible en [https://www.inei.gob.pe/media/MenuRecursivo/publicaciones\\_digitales/Est/Lib1154/libro.pdf](https://www.inei.gob.pe/media/MenuRecursivo/publicaciones_digitales/Est/Lib1154/libro.pdf)

Jones, A., Koolman, X., & Rice, N. (2006). Health Related Non Response in The British Household Panel Survey and European Community Household Panel: Using Inverse Probability Weighted Estimators in NonLinear Models. Journal of The Royal Statistical Society. Series A (Statistics in Society), Vol. 169, No. 3.

Kish, L. (1963). Changing strata and selection probabilities. Proceedings of the Social Statistics Section, American Statistical Association, 124-131.

OIT (Organización Internacional del Trabajo) (1972), Employment, incomes and equality. A strategy for increasing productive employment in Kenya. Ginebra, OIT, enero.

OIT (Organización Internacional del Trabajo) (2002) “El trabajo decente y la economía informal”, 90ª reunión de la Conferencia Internacional del Trabajo (CIT), Informe VI. <http://www.ilo.org/public/spanish/standards/relm/ilc/ilc90/pdf/rep-vi.pdf>

OIT (2003), Informe de la Conferencia, 17º Conferencia Internacional de Estadísticos del Trabajo. Ginebra. Recuperado de: [http://www.ilo.org/wcmsp5/groups/public/---dgreports/--stat/documents/meetingdocument/wcms\\_087590.pdf](http://www.ilo.org/wcmsp5/groups/public/---dgreports/--stat/documents/meetingdocument/wcms_087590.pdf)

OIT (Organización Internacional del Trabajo) (2013), La medición de la informalidad: Manual estadístico sobre el sector informal y el empleo informal, octubre. Naciones Unidas, Fondo Monetario Internacional, Banco Mundial, Comisión de las Comunidades Europeas y Organización para la Cooperación y Desarrollo Económicos (2009), Sistema de Cuentas 2008, Nueva York.

OIT (2013), Resolución I, Resolución sobre las estadísticas del trabajo, la ocupación y la subutilización de la fuerza de trabajo de la 19º Conferencia Internacional de Estadísticos del Trabajo. Ginebra. Recuperado de: [http://www.ilo.org/wcmsp5/groups/public/---dgreports/--stat/documents/normativeinstrument/wcms\\_234036.pdf](http://www.ilo.org/wcmsp5/groups/public/---dgreports/--stat/documents/normativeinstrument/wcms_234036.pdf)

OIT (2018), Resolución sobre las estadísticas de las relaciones de trabajo de la 20º Conferencia Internacional de Estadísticos del Trabajo. Ginebra. Recuperado de: [https://www.ilo.org/wcmsp5/groups/public/---dgreports/--stat/documents/meetingdocument/wcms\\_619087.pdf](https://www.ilo.org/wcmsp5/groups/public/---dgreports/--stat/documents/meetingdocument/wcms_619087.pdf)

ONU. (2009). Diseño de muestras para encuestas de hogares. Directrices prácticas. Recuperado el 15 de Noviembre de 2017, de [https://unstats.un.org/unsd/publication/seriesf/Seriesf\\_98s.pdf](https://unstats.un.org/unsd/publication/seriesf/Seriesf_98s.pdf)

ONU et al. (2016). Sistema de Cuentas Nacionales 2008. New York: Naciones Unidas, Comisión Europea, OCDE, FMI, Banco Mundial. Disponible en <https://unstats.un.org/unsd/nationalaccount/docs/SNA2008Spanish.pdf>

Valliant, R., Dever, J. A., & Kreuter, F. (2013). Practical Tools for Designing and Weighting Survey Samples. New York: Springer.

Willis, Gordon (2004). Cognitive Interviewing: A Tool for Improving Questionnaire Design.

## 11. ANEXOS

### 11.1. Anexo 1: análisis del número óptimo de donantes

En este anexo se presenta una evaluación de la metodología de los donantes por medianas condicionadas utilizada en la Encuesta Suplementaria de Ingresos. El objetivo de este análisis es evaluar el número de donantes óptimos para realizar esta donación. Esto debido a que gran parte de los casos donados tienen sólo un donante.

Esto se comprobará a través de un ejercicio que consiste en excluir la información de ganancias<sup>45</sup> del 20% de los informantes por cada sector de actividad económica agregado, simulando ausencia de información y utilizando lo declarado para el contraste de los resultados. El método más robusto será aquel en el que esté presente un menor error cuadrático medio entre el valor estimado y el verdadero valor entregado por el informante. Para hacer más robustos los resultados, se realizaron 100 simulaciones.

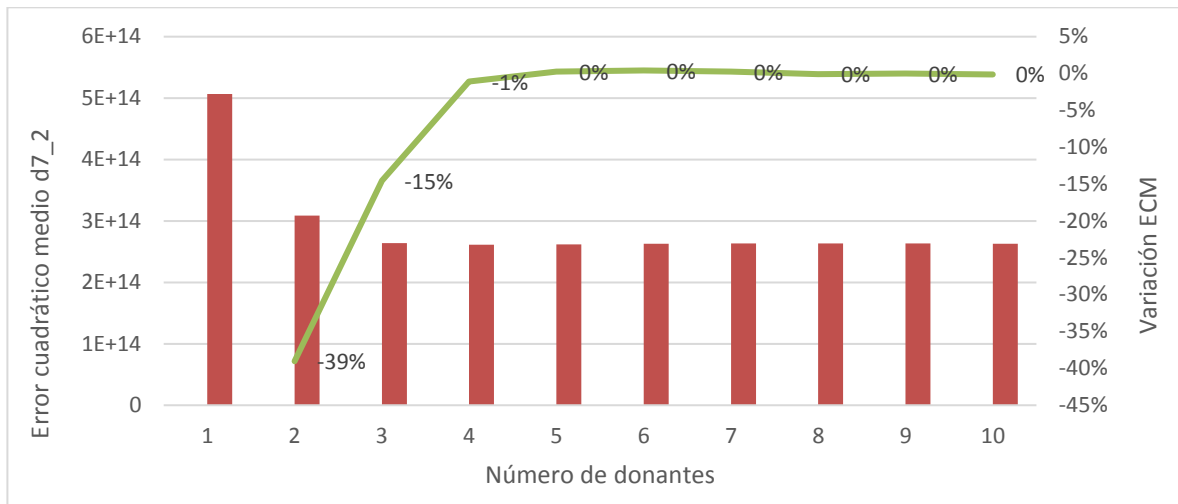
Considerando el promedio del Error cuadrático medio por número de donantes, se puede apreciar que este disminuye 39% al aumentar el número de donantes a 2. Además, si se aumenta el número de donantes a 3 aumenta 15% adicional. Luego de este número se observa que no existen variaciones importantes en el error cuadrático medio.

Se puede observar que estos resultados también son aplicables en la Encuesta suplementaria de ingresos. Considerando los perfiles de donantes utilizados en la ESI se puede apreciar que aumentar el número de donantes reduce el error cuadrático medio mejorando la predicción del valor verdadero, tanto para trabajadores dependientes como independientes.

---

<sup>45</sup> Para fines de este ejercicio se considerará la variable de ganancia declarada (d7\_2), no obstante, es importante indicar que los resultados de este ejercicio son también aplicables en los estimadores de ganancia estimada y ganancia estimada depurada.

**Gráfico 4: Error cuadrático medio según número de donantes**

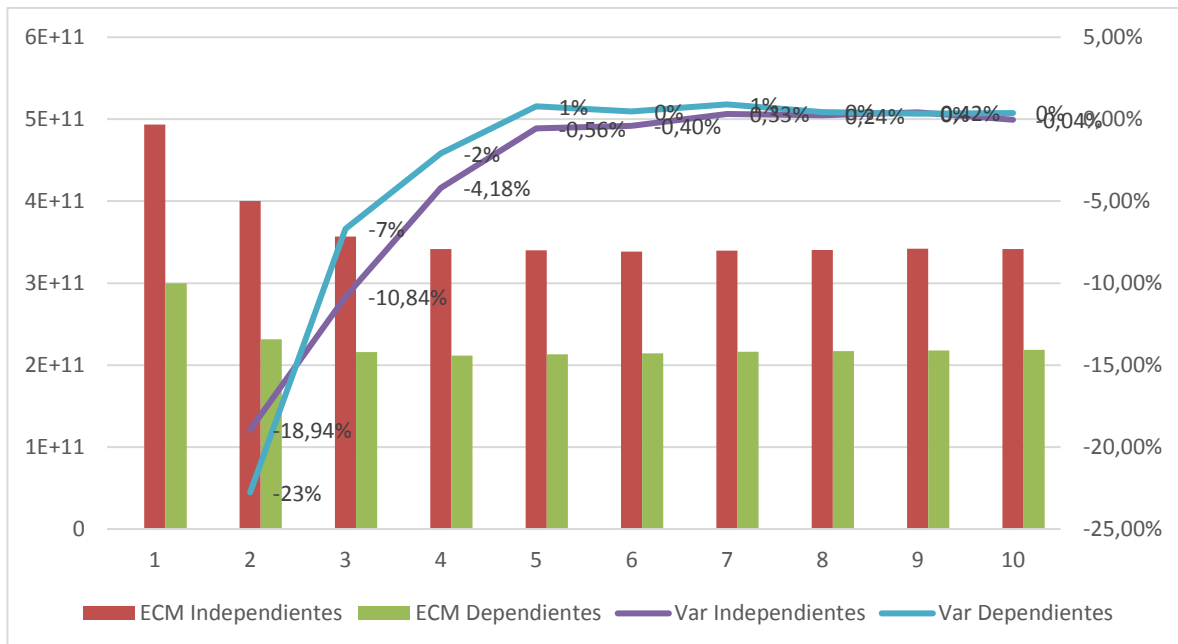


Fuente: Elaboración propia. VI EME.

En el caso de los trabajadores independientes se muestra una reducción de 23% al aumentar el número de donantes a dos, de 7% al aumentar el número de donantes a 3 y de 2% al aumentar el número de donantes a 4. Por sobre este número no se aprecian reducciones estables del Error cuadrático medio.

En el caso de los trabajadores independientes en la ESI, se replican estos resultados. Se observa una disminución de 18,9% al aumentar a 2 el número de donantes, 10 adicional al aumentar a 3, y de 4% al incrementar a 4. Sin embargo, aumentar el número de donantes a más de 4 no provoca mejora en la predicción de los datos.

**Gráfico 5: Error cuadrático medio según número de donantes**



A la luz de estos resultados en la VI EME decidimos aumentar el mínimo de donantes a tres para mejorar los niveles de predicción de la Encuesta. No se consideró incrementar el número a 4, debido a que en la EME no se observó una disminución del error cuadrático y en la ESI la reducción del error al aumentar de 3 a 4 donantes fue baja.

## 11.2. Anexo 2: perfiles de donantes

NIVELES DE DONANTES GANACIAS (N° DE CATEGORÍAS POR VARIABLE)												
1	2	3	4	5	6	7	8	9	10	11	12	13
Actividad económica (18)	Actividad económica (18)	Actividad económica (18)	Actividad económica(8)	Actividad económica(8)	Actividad económica(8)	Actividad económica(8)	Actividad económica(8)	Actividad económica(8)	Actividad económica(8)	Actividad económica(8)	Actividad económica(8)	Actividad económica(8)
Sexo(2)	Sexo(2)	Sexo(2)	Sexo(2)	Sexo(2)	Sexo(2)	Sexo(2)	Sexo(2)	Sexo(2)	Sexo(2)	Sexo(2)		
Grupo de edad(11)	Grupo de edad(6)	Grupo de edad(6)	Grupo de edad(6)	Grupo de edad(5)	Grupo de edad(5)	Grupo de edad(5)	Grupo de edad(5)	Grupo de edad(5)				
Nivel educación(6)	Nivel educación(6)	Nivel educación(6)	Nivel educación(6)	Nivel educación(6)	Nivel educación(6)	Nivel educación(3)	Nivel educación(3)	Nivel educación(3)	Nivel educación(3)	Nivel educación(3)	Nivel educación(3)	Nivel educación(3)
Jornada de trabajo(16)	Jornada de trabajo(16)	Jornada de trabajo(9)	Jornada de trabajo(9)	Jornada de trabajo(9)	Jornada de trabajo(9)	Jornada de trabajo(7)	Jornada de trabajo(7)					
CISE(2)	CISE(2)	CISE(2)	CISE(2)	CISE(2)	CISE(2)	CISE(2)	CISE(2)	CISE(2)	CISE(2)			
Tamaño de la Unidad económica(5)	Tamaño de la Unidad económica(5)	Tamaño de la Unidad económica(5)	Tamaño de la Unidad económica(4)	Tamaño de la Unidad económica(4)	Tamaño de la Unidad económica(4)	Tamaño de la Unidad económica(4)	Tamaño de la Unidad económica(4)	Tamaño de la Unidad económica(3)	Tamaño de la Unidad económica(3)			
Inscripción SII(7)	Inscripción SII(7)	Inscripción SII(4)	Inscripción SII(4)	Inscripción SII(4)	Inscripción SII(3)	Inscripción SII(3)						
Presencia Contabilidad(6)	Presencia Contabilidad(4)	Presencia Contabilidad(4)	Presencia Contabilidad(4)	Presencia Contabilidad(3)								

NIVELES GEOGRÁFICOS DE DONANTES GANACIAS (N° DE CATEGORÍAS POR VARIABLE)				
1	2	3	4	5
Manzana o Sección (2.109)	Estrato (158)	Provincia (50)	Región (15)	País(1)